

nature

5 January 1978

UNCSTD—not just for bureaucrats

In mid-1979 the United Nations will be convening a conference on science and technology for development (UNCSTD), probably in Vienna. It will be a large affair and can be seen as a successor to, and in some ways a consolidation of the earlier conferences on environment, population, food, women, habitat, water and deserts. Indeed those with long memories may recall that in 1963 the United Nations convened a similar large assembly in Geneva on 'the application of science and technology for the benefit of the less-developed areas'—a conference which left a very modest heritage. Many are wondering whether the 1979 conference will be any better than yet another forum for high-flown sentiments, thinly veiled abuse and meaningless resolutions. What happens in 1978, and particularly what scientists do for themselves will help to decide whether the meeting rises above the level of a circus.

UNCSTD is not about science and technology *per se*, it is about national and international policy on these matters. There are four expressed main objectives:

- to decide on ways and means to apply science and technology in establishing a new international economic order,
- to strengthen the technological capacity of developing countries to enable them to apply science and technology to their own development,
- to realise scientific and technological potential in solving development problems of national, regional and global significance, especially for the benefit of developing countries, and
- to provide instruments of cooperation to developing countries in the use of science and technology for solving socio-economic problems that cannot be solved by individual action.

Talk of a 'new international economic order' does not, of course, exactly commend the conference to those developed nations which are suspicious of this concept;

even so, they, along with developing nations, have received a detailed request for the submission of national papers relevant to the conference and it is from these papers, together with the proceedings of regional conferences now being held, that the detailed agenda for UNCSTD will be forged.

The value of such a conference is undoubtedly that it causes some important people in government to have to stop and think about policy or lack of it. In this way the application of science and technology for development might no longer be regarded as a bit of an insignificant backwater in the developed world or as something which has to undergo violent changes in its level of support in the developing world.

The weakness of UNCSTD, however—or at least its potential weakness—is that scientists and technologists may not get a look in; the conference could be entirely bureaucrats talking to bureaucrats. It is possible for this to happen; the conference's secretary-general, Joao da Costa, a Brazilian scientific diplomat, fully admits in a recent issue of *Development Forum* that the secretariat has no authority to bypass governments and go directly to the masses—all it can do is encourage governments to hold discussions with scientists, trade unions, professional organisations and so on. Thus the voice of the practitioner need only be heard if the government chooses to ask for it.

There is, therefore, a need this coming year for scientists and technologists to speak up and use every channel that is open to them to tell their colleagues, in their own countries and around the world, how science and technology is being used and how they think it should be used for development. This was an appeal made very clearly and eloquently in *Nature's* columns recently by Michael Moravcsik of the University of Oregon (24 November, page 228) and we hope that his message will encourage more people to toss their ideas into the arena. Certainly *Nature* will try to be receptive over this coming year to those in the developing world who wish to speak to a large audience. □



Out of the frying-pan into the PWR

Amory B. Lovins, a consultant physicist, and British Representative of Friends of the Earth Inc., the US sister-group of the independent UK group FOE Ltd., argues that US pressurised water reactors are an economic disaster

THE juggernaut advance of light-water reactors (LWRs), and especially pressurised-water reactors (PWRs), through world nuclear markets has seemed to their promoters a sort of manifest destiny. When Britain rejected PWRs for the second time, in 1974, the president of Westinghouse, which first developed them, replied magisterially that this was only "an interim solution, as we are not convinced that it addresses Britain's long-range energy needs." Since then, British electricity use has risen by less than 2%, while excess output capacity above a reasonable 20% reserve margin has risen to 20% of peak output and may soon be 30 to 45%. Closer analysis of energy use has also revealed not only that far more efficient use and even proven renewable sources are practicable and advantageous, but also that Britain already has twice as much electricity as is needed for the premium uses that can give value for money from this very costly, high-quality form of energy. Yet Westinghouse and its allies—including European vendors, the National Nuclear Corporation (NNC), and the Central Electricity Generating Board (CEGB)—are now back again. They argue that the British preference for advanced gas-cooled reactors—the last major indigenous programme of thermal reactors outside Canada—is a sentimental anomaly that now, bereft of further excuses, must give way to technical and economic rationality.

In an open letter to the Prime Minister of the United Kingdom, James Callaghan, on 21 December, Friends of the Earth document a counter-argument. They suggest that the worldwide dominance of PWRs has arisen not from merit but from aggressive salesmanship and "a process of 'mutual intoxication' whereby US promotional institutions persuaded each other, then their counterparts abroad, that LWRs' supposed merits were real and had been demonstrated when in fact they had not and still have not been realised: the distinction between promotional prospectus and critical evaluation was completely obscured". Thus, according to a forthcoming study*, the gas-graphite system displaced in France by PWRs was in retrospect superior, but policy-makers were caught up in a skilfully propagated and wholly unfounded PWR euphoria. Systematically mistaking hope for fact, they thought they *knew* how much PWRs would cost to build and to run, how reliable they would be, and what fuel burnup they would attain. US experience has fallen far short of these expectations. Real capital costs have averaged more than twice as high as promised; real fuel-cycle costs will be about six times as high. Capacity factor has averaged 20 to 30% lower and burnup 20 to 40% lower than forecast. The picture is not getting brighter. Two recent studies, for example, show PWR capital costs rising by 20% y^{-1} or \$188 kW(e) $^{-1} y^{-1}$ in constant 1976 dollars—about three times as fast as coal-fired stations. Much European and Japanese experience is similarly disappointing. Yet NNC's latest report on reactor choice is not an analysis of independently established facts so much as a seller's advertisement of alleged virtues. It is remarkably like the promotional paper that USAEC staffers drafted for the European Community's "Three Wise Men" when they recommended twenty years ago that Europe switch to LWRs.

In the US, the euphoria has worn off abruptly. The doubtful economics of real (as opposed to paper) PWRs, uncertainties about demand, reliability, and safety, and the macroeconomic problems of a capital intensity 10 to 30 times that of new North Sea oil capacity have together led to what FOE call "the most dramatic collapse of a major industrial enterprise in history". US yearly domestic orders in 1972–6 (net of cancellations ascribed to the year of original ordering) were respectively 28, 38, 17, 5, and 3 reactors. In 1975–6, seven reactors were ordered, eight cancelled, and 13 deferred (six indefinitely). In 1977, two were ordered, four cancelled, and 26 deferred. This quickening disintegration has led Dr Schlesinger's deputy to state that "the nuclear option has essentially disappeared" in the US.

Official forecasts of US nuclear capacity in the year 2000 have been falling so quickly that, extrapolating linearly, the 1978 forecast for 2000 should be zero (in fact the asymptote might be as high as 2½% of present US delivered energy use, or about the level of firewood). Nuclear expectations are likewise plunging in France, Germany, the UK, Japan, and elsewhere at such a rate that the 1979 forecast of 1985 OECD nuclear capacity should be zero. In Canada, which has had none of the US regulatory problems, forecasts have dived just as steeply as in the US, suggesting that the cause is not some political artifact but fundamental market forces—which President Carter's energy policy does nothing to discourage and much to reinforce. In short, PWRs are proving all but unsaleable throughout the industrial world. Staunching the rapid haemorrhage of money and staff from reactor vendors is requiring proliferative exports to developing countries, lavishly subsidised by exporting governments. This is not much of a vote of confidence in PWRs or in Britain prospects for profitably exporting them. Indeed, reports in the financial press suggest that no LWR vendor has sold reactors at a profit: vendors' cumulative losses are said to exceed £1 000 million in the US (over half of it for Westinghouse alone), £250 million in Japan, over £200 million in West Germany, and large but undisclosed sums in France and elsewhere. This hardly seems a promising line for Britain to follow—salvaging and refloating other countries' lame ducks.

Further, while PWR advocates consider the vexed question of pressure-vessel rupture adequately resolved, FOE (like most US observers) have long put this issue rather far down a very long list of serious safety problems. Their letter documents ten compendia showing over 200 unresolved major safety problems of PWRs, most of them officially acknowledged. British advocates' optimism about PWR safety may rest not on the detailed knowledge which they claim, but on the lack of it which they have in the past displayed (notably in the 1973–4 controversy) and which the uncharitable can infer from the 1977 NNC report. It is possible that the NNC, CEGB, DOE, and Nuclear Inspectorate staff have in fact all done their homework. There is, however, no published reason to believe this is the case. They may still not have consulted the main original sources, preferring to rely, as in 1973–4, on summaries prepared by US parties that could hardly be considered disinterested. FOE have therefore proposed that specified papers well known in the US debate be the subject of public colloquy in Britain between highly qualified US critics and their British adversaries.

Only four years ago, the NNC and CEGB were pressing for an urgent programme of 32 PWRs—which, if adopted, might well by now have pushed the British electronuclear industry over the brink of ruin. The advocates of haste at that time were wrong; the advocates of caution were right. So it may be again. □

*Light Water: How the Nuclear Dream Dissolved (New York: Basic Books, March 1978), by Professor I. C. Bupp (Harvard Business School) and Dr J.-C. Derian (an official of the French regional development ministry DATAR).

German leader for JET project

Dr Hans-Otto Wüster, a West German accelerator physicist presently on the directorate of CERN, was named on 20 December as project leader for the building of JET, Europe's planned giant torus for extending fusion energy research. Dr Paul Rebut, the French physicist who has led the design team for JET, has been named deputy, and Dr Romano Toschi (currently head of the Italian fusion effort based at Frascati) will be Chairman of the Management Committee. All three appointments are provisional, but are expected to be adopted formally by the Council of Ministers early in 1978.

Dr Wüster's appointment is bound to be felt as somewhat of a rebuff to Rebut; but it was not unexpected. This is not because of any failing in Rebut, but because of the vexed European politics that have dogged JET from its inception. When the decision was made last year that JET should come to the Culham fusion laboratory in the UK (where Rebut's design team had been based) rather than to Garching in West Germany, it was felt that some kind of recompense should be paid to Germany. The appointment of Wüster to the top job achieves this—if to the chagrin of France.

Wüster is not well known at Culham—his field has been high energy physics, not fusion—and he

may find his initial reception cool. But he has time yet to work out his relationship with Rebut and Toschi and those of the design team still left at Culham. The Interim JET Council that appointed him, composed of representatives of all the nations contributing to JET, is suffering from the usual problems of large international committees and is not making its decisions particularly rapidly. It seems unlikely that the team will be completed before July, and meanwhile the main competitor—The Princeton Tokamak—has a chance to forge ahead. After all the earlier delays in the JET project Culham staff look on this further constraint somewhat resignedly.

Wüster himself is no stranger to large international scientific projects. He was deputy to John Adams during the building of the 400 GeV SPS accelerator. Once it was agreed to build it, the SPS became a great European success—built under budget, ahead of schedule, and beyond specifications. Wüster managed the budget of the SPS, and no doubt his success there influenced the JET Council. It will be interesting to see if he can be successful with the budget and management of JET, in a technology he knows less well.

Staff at CERN respect Wüster as a man with great administrative ability, a man who has his goals very clearly



Hans-Otto Wüster: goals clearly in mind

in mind, a blustering, jovial man "who will pile in rapidly" and take to his job with enthusiasm.

Another undoubted asset Wüster will bring to JET is his long-standing friendship with his opposite number at Princeton: Dr Paul Reardon who heads the Princeton Tokamak team. Reardon was also a high energy physicist; and he played a similar role to Wüster's in the building of the Fermilab National Accelerator, the 400 GeV US competitor to Europe's SPS. Fermilab was built first, and skimmed the cream, because of the delays that hit CERN in finding a site and funding in Europe. Wüster will no doubt go all out to prove that Europe can, sometimes, be first.

Robert Walgate

Construction shares

ONE of the most pressing tasks the UK research councils will have to tackle early in the new year will be to decide how to spend their individual allocations of the money promised them late last November and early December for construction and energy conservation. Firm decisions must be made very soon because the money has to be spent in 1978-79.

The Advisory Board for the Research Councils (ABRC) announced its decision on how to divide the total sum between the individual research councils just before Christmas: a total of £4.5 million had to be shared between four research councils (the Social Science Research Council did not put in a bid) and the British Museum (Natural History). Initially, £4 million had been set aside for the recipient bodies of the science budget to spend on construction during 1978-79; but in December, when the Secretary of State for Energy announced the UK's £320 million energy conservation programme, the sum was increased to £4.5 million. £1.1 million of it is to

be spent on energy conservation.

The £4.5 million has been divided as follows: the Agricultural Research Council is to receive £1.07 million, of which £0.27 million is for energy conservation; the Medical Research Council, £0.73 million, of which £0.13 million is for energy conservation; the Natural Environment Research Council, £0.65 million, of which £0.15 million is for energy conservation; the Science Research Council £1.75 million; and the British Museum (Natural History) £0.3 million.

The Natural History Museum probably has the firmest plans of what to do with its share. Most of it will be used to provide accommodation for fossil collections and staff at the museum's station at South Ruislip whilst preparations for a new exhibition hall in part of the main building get underway.

As yet, however, the research councils claim to have few firm plans for spending their shares, even though they must each have put forward definite proposals to the ABRC. But they do agree that major new capital projects are out of the question:

the fund is too small and the time available too short. Most of the money will be spent on extending or improving existing buildings or speeding up construction work already started. For example, the SRC will use some of its share to speed up work on the nuclear structure facility at the Daresbury Laboratory and to extend some of its central user facilities. The NERC may use some of its share to put into operation existing plans to extend facilities at the Institute of Hydrology.

Judy Redfearn

BBC commemorates Bronowski

GEORGE Steiner, a controversial man of letters and champion of science, who learned physics from Enrico Fermi and chemistry from Harold Urey, will broadcast live in the UK on BBC2 television next Tuesday at 8.10 pm in the first Jacob Bronowski memorial lecture. Bronowski, devoted to what he called 'the democracy of the intellect', literally gave himself to an epic tele-

vision series—'The Ascent of Man'—which presented his personal and heroic view of science as the saviour of mankind. Bronowski died shortly after completing the programme; these lectures are the BBC's tribute to him.

Steiner is an interesting choice to lead off the memorial lectures. He has many points of contact with Bronowski; and as many differences. Both were Jewish exiles from Europe: the Nazi holocaust is writ large in both their lives, and both have shown themselves aching to find escape from it. The ultimate intellectual sanctity of science, with its references to verifiable fact, particularly in the physical sciences and mathematics, appealed deeply to both men as the antithesis of dogmatism. But whereas Bronowski searched for a release, in the fashion of The Enlightenment, for all of mankind, Steiner appears to look only towards an elite, a bright spark in the

darkening eclipse of humanity.

Steiner will argue in his lecture that it is pointless to look for controls on the pursuit of science; for knowledge will out, somewhere, sometime. One day scientists will make their test-tube babies. One day there will be genetic engineering of human beings. We are powerless against the march of truth. Steiner will prove himself the hawk of science, where Bronowski was the dove. In the process he may do science a disservice. He is a great acquirer of knowledge: 'What else is there?' he asks. Bronowski also asked for compassion, an emotion that Steiner might be suspected of dismissing as romantic.

In the 'Ascent of Man' Bronowski drew attention to a man who he said was the most intelligent he'd known: Johnny von Neumann, the originator of games theory. Von Neumann failed mankind, said Bronowski, as an intellectual elitist who believed

not in the democracy but the aristocracy of the intellect. Steiner appears to fall into the same mould as von Neumann.

Steiner learned science at the University of Chicago in the late 1940s. He "scraped" a first and to his lasting regret was turned away from science by his careers advisor who said he could not cope with physics (but that was Chicago in the days of Fermi; Lee and Yang were students) and that the rest of science was "bottle-washing". Steiner, fluent in four languages, turned to journalism and ultimately to the study of language.

He would have liked to have been a biologist. His passion is mathematics and his heroes are mathematicians. The creation of melody, and of proofs (in mathematics) that the truth of a certain proposition is undecidable, are, he believes, the greatest mysteries.

Robert Walgate

ESA names European candidates for first Spacelab mission

THE European Space Agency (ESA) has named the four candidates from whom one will be selected to serve, along with one American scientist, as a "payload specialist" on the joint US—European Spacelab mission scheduled for the second half of 1980. The candidates were presented at a press conference on 22 December at ESA's headquarters in Paris. They are Franco Malerba, 31, an Italian electronics engineer and physicist; Ulf Merbold, 36, a West German physicist and staff member of the Max Planck Institut für Metallforschung in Stuttgart, where he has been researching crystal lattice defects; Claude Nicollier, 33, a Swiss astronomer; and Wubbo Ockels, 31, a Dutch physicist who is doing research at the Nuclear Physics Accelerator Institute in Groningen.

With the exception of Malerba, all are married with one child. Ockels is the only non-flier in the group. Merbold is a glider pilot, Malerba holds a private pilot licence and Nicollier, after working as a professional pilot for Swiss Air, is still in the Swiss Air Force.

After three months of Spacelab familiarisation at European facilities and further evaluation in Houston, one of the four candidates will be eliminated. After next May the other three will continue training until a few months before the flight when one of them will be chosen to be the first European to travel and work in earth orbit. The other two will act as back-up specialists on the ground. On the first ESA/NASA mission lasting one week, the two scientists will be responsible for a payload of 76 experiments in the fields of materials science, atmospheric physics, life sciences, solar physics,

astronomy, earth observation and space technology.

"We were not looking for a superman or superwoman," said Michel Bignier, Director of the Spacelab programme. He explained that the scientists would be working in a pressurised "shirt-sleeve" atmosphere. Nonetheless, the evaluation criteria had been particularly severe—equivalent to level II criteria used by NASA to select permanent mission specialists. "Nobody wanted to run the risk of a rejection by NASA," added Mr Bignier.

By last September, 53 candidates had been preselected from 2,000, including 35 women, who had responded to ESA's announcement of opportunity issued in April. One woman, Anny-Chantal Levasseur—regourd of France, was among 12 finalists but failed on the last medical test.

Candidates were screened by four panels of experts covering general

assessment and system engineering, scientific, psychological, and medical tests. All the candidates agreed that the psychological testing, carried out at the DFVLR Institute for Flight Medicine in Hamburg, was particularly arduous. Klaus-Martin Goeters, a member of the institute and head of the psychological panel said that tests were aimed at ensuring that those selected could cope with the stress created by the heavy Spacelab workload and the unusual environment. Drawing a typical personality profile, Dr Goeters described the successful candidate as "someone who is highly motivated, with a tendency to be an introvert, preferring small groups to a larger social context, who is unpretentious in habits and life style, thus capable of envisaging certain privations, and is highly mobile: that is a person interested in and looking for new things—with, you might even say, a taste for adventure".

Betty Werther

Sorry, for copyright reasons some images on this page may not be available online

Claude Nicollier, Ulf Merbold, Wubbo Ockels, Franco Malerba: Europe's spaceman scientists

GMAG: Stormy weather ahead?

1978 may prove to be a difficult year for Britain's Genetic Manipulation Advisory Group. David Dickson reports

GMAG, which was set up by the UK Government following concern at the implications of research into the genetic manipulation of micro-organisms, celebrated its first birthday last month, thereby reaching the half-way stage of its first term of office. So far both scientists and trade unionists have given it, in general, a clean bill of health; comments on its performance range from "as good as could be expected" through "better than anticipated" to "a highly successful experiment"

Applications for more than 80 experiments, from about 25 different institutions, have been received, and all have been approved. According to Sir Gordon Wolstenholme, director of the CIBA Foundation and chairman of GMAG, the first year has been "surprisingly profitable and useful".

Yet if the first year has been relatively calm, there are storm clouds on the horizon. Two factors in particular are likely to put pressure on GMAG's activities in the near future, and to test its skill in treading a delicate path between different—and sometimes opposing—interest groups.

The first of these factors is industry's rapidly growing awareness of the financial rewards offered by genetic engineering techniques, from the production of artificial hormones to, possibly, new plant crops. Already seven major US companies have established recombinant DNA research laboratories, and several British companies, after a relatively slow start, are jumping on the band-wagon.

The second factor is the additional knowledge, both about genetic structures and the potential hazards of their manipulation, that has emerged in the four years since a temporary moratorium was first suggested by US scientists. This knowledge suggests that some of the initial reaction was over-cautious, and that current guidelines might be "safely" relaxed.

Of these two, it is the industrial applications that have so far impinged most directly on GMAG's activities, primarily through the issues of confidentiality. For genetic engineering is rapidly becoming a field in which the dividing line between science and commerce is frequently indistinguishable.

In the US, for example, shortly after a team of San Francisco scientists headed by Dr Herbert Boyer confirmed that it had successfully inserted into bacteria a gene capable of producing the hormone somatostatin, a small Californian company Genentech an-

nounced that it would be using the technique to produce somatostatin for sale to research organisations and pharmaceutical companies by mid-summer 1978, at a cost considerably below the current market price.

In Britain, the National Research and Development Corporation—the body responsible for patenting and exploiting the results of government-sponsored research—recently held a lively and well-attended meeting for university scientists on the potential applications of genetic engineering. NRDC has already awarded a research contract to the John Innes Institute near Norwich for the study of improvements in the production of the antibiotic streptomycin, and other projects are in the pipeline.

Although some companies, such as ICI, have been studying the applications of genetic engineering techniques for a number of years, little is publicly known about how far such work has progressed. For with the potential rewards so enormous, not only commercial companies, but all those with potentially-patentable techniques, are keen to play their cards as close as possible.

There is therefore an inevitable reluctance to reveal even in outline details of future research programmes to a group such as GMAG over whose use of the information—notwithstanding the Official Secrets Act—an individual company can exercise no direct sanction.

The confidentiality issue is one which the difficulty of reconciling "health and safety" factors against "national interest" has provided GMAG with some uncomfortable moments. In particular, the four trade union representatives—two from the Association of Scientific, Technical and Managerial Staffs (ASTMS), one from the Institute of

Professional Civil Servants, and the medical officer of the Trade Union Congress—while respecting the need to maintain confidentiality for commercial reasons, have refused to be bound by any secrecy agreement which might prevent them from serving the interests of their members.

Many hours of discussion have been spent trying to devise a formula by which secrecy can be guaranteed for research programmes which must soon, under new health and safety legislation, be reported to the Health and Safety Commission. Eventually a compromise formula has been agreed under which, for a trial period of six months, members of GMAG will agree to a simple secrecy pledge, and applications for commercially-sensitive experiments will be seen by all except those considered to have some direct interest.

But for most scientists, the confidentiality issue, apart from interfering with claims for scientific precedence, remains largely peripheral. Of far greater concern is the type of pressure that may be put on those responsible for maintaining safety guidelines as a result of industry's interest.

In particular, many fear that industry's desire to pursue and exploit the applications of such research as quickly as possible may result either in pressure on governments to reduce containment requirements too far, or to remove them altogether, leaving the management of risk to the vagaries of the market place (and the accountancy skills of insurance companies).

There are already pressures from scientists themselves to rationalise present containment levels, bringing them more in line with procedures adopted for other types of experiment using known pathogens, and taking into account both the development of new genetic engineering techniques and of more precise assessments of hazards.

Research by Professor Roy Curtiss at the University of Alabama in the US, for example, has confirmed that the K-12 strain of *Escherichia coli*

Sorry, for
copyright reasons
some images on
this page may not
be available online

Above: view of *E. coli* by scanning electron microscope;
right: researchers working in a high containment laboratory



NH

most commonly used in present recombinant DNA research is unable to survive in the human intestine like normal strains, and is thus very unlikely to cause an epidemic if it escaped from the laboratory.

This research has also led to the development of a "disabled" strain X-1776 which has already been passed by GMAG for some high-risk experiments. And it is widely expected that the successful development of a host-vector system using another "disabled" strain of *E. coli* with equally "safe" properties will be announced soon in the UK.

Such developments are already encouraging a certain relaxation of physical containment levels, in particular since the greater the biological safety of the organism, the lower the necessary degree of containment. In the US, for example, the National Institutes of Health published last summer proposed revisions to the guidelines first issued in June 1976. The new draft guidelines, which were publicly debated in Washington just before Christmas, and have already aroused considerable opposition from environmentalist groups such as Friends of the Earth, involve a considerable reduction in the containment levels for certain widely-used experiments.

Perhaps the most controversial proposal is to reduce the required physical containment level for experiments using recombinant DNA from mammals other than primates from P3 to P2; in practice this means that experiments using, say, mouse DNA which under the current regulations require purpose-built facilities, could be carried out on an open laboratory bench with a few relatively minor adaptations.

At present it is uncertain whether this particular proposal will be accepted (although others, such as a revised system defining physical containment requirements, probably will). Certainly many British scientists feel that, in the present state of knowledge, this may be relaxing things a little too far, at least for "shot-gun" experiments using unpurified DNA.

Yet the main concern about proposed revised guidelines is not so much their scientific content, but the fact that, if adopted, serious discrepancy between British and American laboratory practices could result, with potentially embarrassing consequences.

It could, for example, be cheaper to send a research worker to the US to carry out a set of experiments requiring little more than a conventionally-equipped laboratory bench, than to instal—or even hire—facilities providing the higher physical containment levels required to carry out the same experiment in the UK.

Such embarrassment is likely to be reinforced if other European countries, most of which have so far followed the British guidelines as suggested by the Williams Committee in 1976, decide to break ranks and, in line with the US, introduce significantly lower containment levels (as the French are, indeed, now proposing to do).

If, for example, Germany decides to introduce less stringent guidelines, British scientists working at the European Molecular Biology Laboratory at Heidelberg, which operates under German laws, may find they are permitted to carry out experiments which they are unable to do at home.

In such a situation, GMAG could find itself caught in a more difficult dilemma than that which it has already faced over confidentiality. For while responding—reasonably—to demands from scientists to rationalise containment levels, public confidence requires that it must not be seen to be responding to outside, and particularly industrial, pressure.

The members of GMAG are conscious that if they impose too harsh a set of restrictions on industrial research programmes as compared with other countries, companies will merely transfer their research programmes—and their revenue-earning potential—elsewhere; "and we don't want a repeat of the penicillin story", according to one GMAG member.

However, the Williams Report, which itself followed the public discussions generated by the Ashby report of 1975 and received evidence from a wide range of bodies, is still less than two years old. Any major attempt to re-draft its proposals at this stage is likely to be seen as premature, if not by scientists at least by the public, to whom some of the more fanciful possibilities (and dangers) of genetic engineering have only recently come home.

Indeed, it is one of GMAG's strengths that, through the trade union and "lay" representatives, it is able to take public concern directly into consideration (thus avoiding the major confrontations between scientists and the public that have, for example, occurred in the US).

Yet as the pressures on GMAG increase, so the tensions between its constituent parts may become more difficult to control. Such tensions arise partly from the ambiguous nature of GMAG itself, seen by some as an unfortunate, if necessary, part of scientific bureaucracy, and by others—including the trade unionists—as a successful model of public participation in research policy, with potential applications in other areas.

Disagreement over how to handle the confidentiality issue has already

brought these tensions to the surface. And while there is little criticism of the behaviour of the trade union representatives on scientific aspects of GMAG's work, they have already been accused of trying to make "political" points out of issues such as the future of the Ministry of Defence's Microbiological Research Establishment at Porton Down (part of which is likely to be transferred to the Department of Health and Social Security for carrying out top risk experiments requiring Category IV containment levels).

The political temperature of the debate has been raised by the recent decision of the British Society for Social Responsibility in Science to set up a genetic engineering group. This has already attacked the "complacent" attitudes of many scientists towards the safety aspects of such work. And the BSSRS group also feels that safety guidelines should have the force of law.

These issues will also undoubtedly be aired at the public hearings into genetic engineering soon to be held by the House of Commons Select Committee on Science and Technology. Reactions to the decision to organise these hearings vary; some are wary of the effects that renewed public exposure will have, and of possible "anti-scientific feeling" on the committee, while the BSSRS group has, in contrast, given the decision enthusiastic support.

But there are a number of important issues in need of attention. One of these is the make-up of GMAG itself, whose members are due to be reappointed at the end of 1978. Many would like to see more scientists on the group, particularly experts in fields of research not adequately covered at the moment. And there is still a residual unhappiness that although GMAG has two ASTMS representatives, there is no member of the Association of University Teachers acting in a representative capacity.

Another issue to which the Select Committee may give its attention is the intensive, quantitative study of the potential hazards of genetic engineering research. So far, despite an apparent commitment both in the US and the UK to the need for such work, little has yet been done by the scientific community actively to promote it.

GMAG is due to present its first annual report to Mrs Shirley Williams, Secretary of State for Education and Science, in a few weeks time. One thing is already clear: that GMAG's role is likely to become both more important and more difficult. As Sir Gordon says: "If anything, our 'political' problems are not going to die out, they will increase; and not because anything is going wrong, but because it is in the nature of the animal." □

correspondence

Austrian science journalists were warned

SIR,—As one of the people principally involved in disclosing the fraudulent activities of Schaden ('Austrian science minister attacks science journalists in fraud case' 24 November, page 292), I should like to point out that in 1973 I warned Austrian scientists and science journalists about the questionable 'discoveries' of Schaden and Celta. Unfortunately, as the Science Ministry's document demonstrates, this was to little avail.

The purpose of the document and the Ministry's previous interventions and warnings, was to protect the scientific community from public discredit on account of a charlatan's 'research'. It was also to serve as a constructive warning to scientific journalists not to take every apparent scientific 'break-through' at its face value.

If the Science Ministry's document lacks thoroughness, this will certainly be supplied by the court's records of the Schaden case.

Yours faithfully,

WILHELM GRIMBURG

Vienna, Austria

A third world energy view

SIR,—Effective use of the world's finite cheap oil reserves to save time and effort deserves a broader discussion than that given by Alvin M. Weinberg (20 October, page 638). Efficiency of the conversion of energy into time saved is highly variable. To travel 15 miles, for large numbers of people, means walking in the tropical heat carrying luggage and taking, say 5 hours. To be lucky enough to have a small motorcycle (and do the journey in 30 minutes on 0.1 gallons of fuel) saves 45 gruelling man hours per gallon. To then change to a large car (and come back at 90 miles per hour using 1 gallon of fuel) will only save a further 0.37 man hours per gallon.

The assertion that "a free society allows each of us to make the choice" and "economics . . . integrates all these . . . judgments" means that the rich have the right to decide. But when availability of cheap oil begins to decrease these judgments will, no doubt, be re-evaluated, and those who make them now thus carry a heavy respon-

sibility. Surely they will be expected to justify their judgments when others reach the position where they could use oil but find that it has all been carelessly squandered to save milliseconds of millionaire's time per gallon.

Yours faithfully,

M. D. MELAMED

Maputo, Mozambique

Desert rainfall

SIR,—J. L. Deaton (24 November, page 294) makes several statements concerning our article (19 May, page 192) with which we disagree. He claims, in particular, that "in no sense is the mean too large". There is, however, a well defined sense in which the mean of a positively skewed distribution (for example, precipitation) is "too large". Specifically, it may be desired that a statistic, to be taken as a measure of central tendency, has the property that either it represents the midpoint of the distribution (that is the median), or that it represents the value of the distribution which is most likely to occur (that is the mode). If the distribution is positively skewed, then the mean is larger than the values satisfying either of these properties.

Deaton also claims that the mean "is quite indicative of how much rain commonly falls in most regions—even Gao and Niamey." But the mean, at least in some sense, is not at all indicative of how much rain "commonly" falls in the Sahel. As an example, consider the Gao August precipitation data used in our article: only 4 of the 35 observations (11%) fall within 10% of the mean; 23 of the 35 observations (66%) fall below the mean; and the mode (or most 'common' value) is 80.0 mm, substantially less than the mean of 100.9 mm.

Deaton states that "there seems to be only a slight tendency for the degree of skewness to increase as the average amount of precipitation decreases". This statement contradicts the results of studies on this issue in the climatological literature (Arnold Court in *Climates of North America*. World Survey of Climatology 11, 212 (Elsevier, Amsterdam 1974)). While it is true that all precipitation distributions are at least slightly positively skewed, there is a marked tendency for the

degree of skewness to increase as the mean precipitation decreases.

Finally, Deaton asks for documentation of the observation that "recent weather tends to influence perceptions more heavily than earlier weather and wet spells more heavily than dry ones". The documentation for this statement is a quote from J. C. Caldwell "Rain-fall Statistics, Droughts, and Desertification in the Sahel," *Desertification*, 84 (Westview Press, Boulder, Colorado 1977).

Yours faithfully,

MICHAEL H. GLANTZ

RICHARD W. KATZ

National Center for
Atmospheric Research
Boulder, Colorado USA

IQ or intelligence type?

SIR,—The controversy over race and IQ is clouded by the lack of culturally neutral testing procedures. Who would we be to say that a community of blacks who tested slightly lower than comparable whites were not immeasurably superior in forms of mental ability not emphasised in the test procedures? A slight mean shift from one community to another would doubtless be given more weight than it deserves, and it would still be possible (if its Gaussian distribution was shallower) for the 'less endowed' community to provide more geniuses than its counterpart. Furthermore, the difference between the two means would certainly be less than that between any two members of either group that you might encounter in the street, but would this be taken into account by those with political motivations? Data may be used to substantiate whatever case and in an area like this the phenomenon could have serious consequences. The "compensatory advantages" Sir Andrew Huxley cites in his address to the British Association annual meeting would soon be seen as racially-discriminative favouritism, and is it really so likely that humane considerations would moderate the debate? We take scientific findings much less seriously than it is popular to imagine: the new evidence in fields such as psychokinesis, spoon-bending, E.S.P., and tobacco smoking has done little to alter the attitudes of those with convictions of convenience that contradict the specialist conclusions, and there are still

eminent scientists who refuse to accept evolutionary theory and rely instead on more metaphysical alternatives. Scientific findings do not alter human belief so readily.

Huxley says that it is a characteristic of research that its outcome is not known in advance. That is so: but it is a presupposition of outcome that often motivates research, which may bias its interpretation, and which makes it possible that results will not alter preconceptions. In the correspondence for the same issue in which Huxley's reply to *Nature's* earlier editorial on his address appears (29 September, page 366), C. J. Robbins subscribes to the widely-held view that science sets out to falsify theories through experimentation. As the pages of *Nature* show, this is rarely true; the commonest motivation in research might well be exactly the converse.

The concept of IQ itself is a simplistic parameter which assumes that cognitive, perceptual, deductive and mathematical abilities go hand in hand, whereas our selection of individuals (for employment, etc.) shows this cannot be the case. We accept that brilliant mathematicians may be absent-minded, that musicians may be hopeless fine artists, scientists poor communicators. Why then do we hear so little about what we might call intelligence type? Until we have evolved realistic codes of criteria for assessment that reconcile mental measurements with the realities of life I believe we should postpone spurious research into racially-determined IQ, and recognise it as being ill-founded and premature.

In my view this is the most objective manner in which one could admit the limitations of contemporary science, and the irrelevance of research findings to those determined to subscribe to their own beliefs; these two are topics that are ripe for research.

Yours faithfully,

BRIAN J. FORD

Science Unit, Cardiff, Wales

Soviet genetics

SIR,—As far as I know, gerontologists have been wary, over the years, of passing opinions on human genetics and on human genetics programmes. I was therefore very interested to read the review by Dr Zhores Medvedev, a noted gerontologist, on the development of human genetics in the USSR since Lysenko ('Soviet genetics: new controversy', 28 July, page 285).

The review bears a title which is frightening for anyone knowing the history of genetics in that country. A new controversy is hardly needed at this stage. The international congress of genetics is due this year in Moscow

and even calling attention to controversy might cause unpredictable side effects. Fortunately, renewed anxiety for the fate of human genetics in the USSR seems premature at this stage; the review gives a brief account of the re-emergence of the discipline, and for the rest is an all-out attack on the retiring director of the Institute of General Genetics of the Academy of Sciences, Dr N. Dubinin.

Of those who reinstated human genetics in the USSR Medvedev writes: "all of them were void of real practical knowledge of human or medical genetics", and one is left to wonder how they happened to have the courage to do it. Furthermore, if those who only had knowledge of drosophila genetics, of radiation, rodents, cytology and theoretical human genetics can be criticised for their contribution what can a gerontologist say about human genetics; in what position is he to pass value judgements?

Dr Medvedev, although he does not say so, must have participated in the discussions, since he says they were peaceful, albeit not very productive. Again one wonders how the new Institute of Medical Genetics was established as an outcome of such a lack of productivity.

One cannot help noticing the amount of negative emphasis put by Dr Medvedev on an action which tended to reconstruct human genetics. It is my opinion that, since these negative emphases diminish the whole operation, he should produce the evidence on which they are based. What, for example, is the evidence that D. K. Beliaev had been appointed in 1976 president of the 1978 Congress? This is puzzling, because Beliaev is general secretary, and Tsitsin (wheat hybrids at Lysenko's time) is president.

The part of the review which is devoted to Dubinin is less important. Dubinin, as others, went through the bitter years of the Soviet geneticists and survived. At present, he seems to have developed his own personal views on the inheritance of human abilities; these views are questionable but since they have been publicly castigated, they are not official views which might endanger those not sharing them. It seems arbitrary to associate them with a controversy which might harm human genetics; he is entitled to his opinions, no less than Medvedev to his own and I to mine. So far as some of Dubinin's work on human genetics is concerned, and so far as I can read, he has recently used erroneous techniques in the study of multivariation in quantitative traits in man. However, it might well be that, as director, he signs work from his institute which might be beyond the capacity of his technical

judgment (for example *Dokl. Akad. Nauk.* **230**, 4: 957-960, 1976).

I believe that the most contradictory aspect of Medvedev's review is that Dubinin is turning against the programme he helped to start a short time ago. I believe that the accusation Medvedev makes, that "Dubinin is working hard to suppress by all possible means the development of genuine research in the field of human genetics in the USSR", and which amounts to public condemnation, is so important, that Medvedev is bound to produce evidence for it. It would be sad for those who have helped in the reconstruction, even for those who have given a minimal contribution, if attempts to create new conflicts hinder development. Human genetics is a hot science to handle: it seems useless to make it even hotter. After all, state and interstate budgets for human genetics can be cut easily, both east and west of Greenwich.

Yours faithfully,

ITALO BARRAI

Universita di Ferrara,
Italy

The Messinian salinity crisis

SIR,—Lines in reply to an anonymous critic (20 October, page 646)

When composing lyric verse

Be it critical or worse

It is wise to be quite certain of one's ground,

And not to call absurd

Quite a harmless little word

Lest its meaning be not simple but profound.

Crises evaporitic

Irritate our nameless critic

Rightly so, had we but meant what he implied,

But salinity increased

Until crisis point was reached

Whereupon the fauna disappeared or died.

Oh, 'tis pity I declare

That whatever we prepare

And however clear the message that we send,

There are always colleagues who

Having little else to do

Criticise us when they do not comprehend.

N.B.—The term 'Salinity Crisis' was, I believe, coined by Ruggieri (Systematics Association Publication No. 7 (eds Adams, C. G. and Ager, D. V.), 283-290 (London, 1967)) and referred to the apparent extinction of the marine faunas of the western Mediterranean in Messinian times.

Your faithfully,

C. G. ADAMS

British Museum (Natural History),
London

news and views

T cell hybrids: shortcut or dead end?

from F. Melchers

THREE years ago everybody believed that normal lymphocytes could not be fused with malignant lymphocytes in culture. Then Köhler and Milstein (*Nature* **256**, 495; 1975) successfully fused an immunoglobulin-secreting myeloma with normal B lymphocytes which secreted antibodies specific for sheep red cells. This encouraged many laboratories to produce fused B cell lines secreting antibodies with a variety of required specificities (see *News and Views* **266**, 495; 1977). These 'hybridomas' (or Köhlstein lines perhaps?) represent a breakthrough in immunological research. Effector functions of clones of particular lymphocytes, in this case specific antibody production by B cells, can be made immortal. Consequently the clones can be grown up in large scale preparations that will allow structural and functional analyses of the clonal products.

It is difficult to say how this success came about after all the failures. Improvements in tissue culture techniques and selection of suitable malignant cell lines for fusion certainly helped. Köhler and Milstein recognised (*Somatic Cell Genet.* **3**, 303; 1977) that malignant and normal lymphocytes used in the fusion had to be at a comparable stage of differentiation. As a simple rule B cells, but not T cells, could fuse with B myelomas.

Once the ice had been broken everybody was eager to try similarly to fuse normal T lymphocytes and their malignant counterparts. This has now been achieved in several laboratories (Goldby *et al.* *Nature* **267**, 707; 1977; Hämmerling *Eur. J. Immun.* **7**, 743; 1977; Köhler *et al.* *Eur. J. Immun.* **7**, 758; 1977). For these fusions a suitable malignant cell line with T lymphocytic markers had to be found. Robert Hyman from the Salk Institute in La Jolla provided it in the T lymphoma BW5147. Fusion between normal T cells and this cell line was demonstrated by the codominant expression of genetic markers which

differ in the two cells, such as histocompatibility antigens, the T-cell specific Thy-1 marker and isomeric forms of the enzyme glucose-6-phosphate isomerase.

The ultimate aim of these fusions, however, is to preserve the effector functions of the parent T lymphocytes and thereby create cell lines with which the different effector functions of T lymphocytes could be elucidated. This, however, has proved less easy than had been anticipated. None of the T-cell hybridomas has so far revealed any effector function expected for normal activated T cells. There might be several reasons for this failure: the genome of the tumour cell might suppress the expression of the effector functions of the normal T cells; the chromosomes of the normal cells carrying genes for these effector functions might be lost after fusion and during growth of the T hybridomas. It is also possible that we are unable at present to fuse normal, activated T lymphocytes expressing a given effector function with T lymphoma cells or to assay for T cell effector functions *in vitro* when these functions are provided by individual clones of T cell hybridomas.

Three major immunological functions are attributed to T lymphocytes: help and suppression of B cell growth and maturation and killing of target cells. All three functions are thought to be initiated through clonal selection of specific T cells by antigen, resulting in the growth and differentiation of clones of antigen-specific T cells.

How, then, are these functions to be detected *in vitro*? First and most important, no single cellular assay exists for any T cell function as it does for immunoglobulin-secreting B cells in the form of Jerne's haemolytic plaque assay. Killer cells are most often monitored by ^{51}Cr -release by labelled target cells. This method has recently been improved so that between 50 and 250 specific killer cells can now be detected in one culture, allowing *in vitro* analysis of the frequencies of target cell-specific killer T cell clones

(Teh *et al.* *J. Immun.* **118**, 1057; 1977). The general cytolytic activity, independent of the antigen-binding step, of killer to target cell can also be measured with a ^{51}Cr -release assay. Killer and target cells in this non-specific assay are brought into close contact by agglutinins which act as a glue (Forman & Möller *J. exp. Med.* **138**, 672; 1973; Bevan *et al.* *Eur. J. Immun.* **6**, 150; 1976). The antigen-nonspecific cytolysis which occurs when the killer and target cells are in close contact, is precisely what seems to make successful fusions between activated killer T cells and a malignant T cell line impossible, since fusion necessarily involves close contact of the two cells to be fused. Köhler and his collaborators (*Eur. J. Immun. op. cit.*) have now shown that the close contact of activated T killer cells with the BW5147 T lymphoma line during the fusion reaction kills the T lymphoma cells and thereby prevents probably all killer cells, regardless of their target specificity, from becoming immortal. There is little hope of this nonspecific killing mechanism ever being inactivated during the fusion reaction, allowing the establishment of T cell hybridomas with killer activity.

In vitro assays for T cell help or suppression of B cells seem even less sensitive than killer assays. Furthermore, there still exists considerable controversy over the mechanism of T cell dependent B cell triggering (at its height in *Transplant. Rev.* **23**, 1975). Depending on which immunologist one talks to, quite different structures and actions are expected from the products of a T cell hybridoma. Many immunologists believe that T cells produce antigen-specific factors, which for a long time were thought to be immunoglobulin. This now seems less likely. Some people think that such factors carry Ia antigen, but others cannot find these Ia antigens on their factor preparations. Many also believe that these antigen-specific T cell factors in some way induce B cell growth and maturation to immunoglobulin secretion and are therefore mitogenic for B cells.

This all argues for soluble products effecting T cell help and suppression, but not everyone is yet convinced that cell-cell contact can be ruled out as the basis of the physiological action of T-B collaboration. Conceptual confusion prevails in a situation in which convincing experiments are still somewhat scarce.

The chances of finding a T-cell hybridoma which produces helper factor for a given antigen are indeed very low—approximately 1 in 10^4 for the antigen sheep red cells (Lefkovits & Waldmann *Immunology* 32, 915; 1977). No wonder then that the few T hybridomas which have been tested so far for such antigen-specific helper functions do not produce them (Köhler *et al. op. cit.*). If these factors are mitogenic for B cells it might be predicted that, at higher concentrations, they could polyclonally activate all B cells which are receptive to the mitogenic principle of the factor regardless of the specificity of these B cells for the hapten. Assays for B cell mitogeni-

city have become very sensitive (Andersson *et al. Cell* 10, 27; 1977) and such assay systems might be useful to detect growth-inducing properties of factors produced by T cell hybridomas, or even by T lymphomas.

However, one may still hope that T cell hybridomas will open up a new way to an understanding of normal T cell effector functions of help and suppression and their molecular mechanisms of action. Binding of antigen to T cell hybridomas or to molecules released by them seems one simple first step. Selection of T lymphoma lines which display nonspecific helper or suppressor activities with normal B cells, may also help—keeping in mind the 'rule' that cells at comparable stages of differentiation might fuse more easily with each other. Beyond that immunologists will have to replace ideology by scientific knowledge, obtained through better assays, so that they can detect what may be present in, on, or released by, T cell hybridomas. □

Maeyer-Guignard *et al. Proc. natn. Acad. Sci. U.S.A.* 69, 1203; 1972). The assay can be made quantitative (Kronenberg & Friedman, *J. gen. Virol.* 27, 225; 1975; Greene *et al.*, this issue of *Nature*, page 81). Second, both mouse and human interferon mRNAs have been translated in cell-free systems, mouse interferon mRNA in the wheat germ system (Thang *et al. Proc. natn. Acad. Sci. U.S.A.* 72, 3975; 1975) and human interferon mRNA in cell-free extracts from mouse ascites cells, rabbit reticulocytes and wheat germ (Pestka *et al. Proc. natn. Acad. Sci. U.S.A.* 72, 3898; 1975; Reynolds *et al. Proc. natn. Acad. Sci. U.S.A.* 72, 4881; 1975; Raj & Pitha *Proc. natn. Acad. Sci. U.S.A.* 74, 1483; 1977). Third, human interferon mRNAs have also been translated after injection into *Xenopus* oocytes (Raj & Pitha, *op. cit.*; Cavalieri *et al. Proc. natn. Acad. Sci. U.S.A.* 74, 3287; 1977; Sehgal *et al. Proc. natn. Acad. Sci. U.S.A.* 74, 3409; 1977), and both the cell-free assay and the oocyte assay can also be made quantitative. The products of these translation systems have been characterised as interferon using species and antigenic specificities, and sometimes also by showing the product to have the expected molecular weight. Use of the *Xenopus* oocytes showed that injection of polyadenylated RNAs from induced fibroblast and lymphoblastoid cells yielded fibroblast and leukocyte interferons respectively, demonstrating that human cells can produce two different translatable mRNAs and probably contain two genes for interferon, coding for polypeptides differing in primary sequence.

What has been learnt about the control of interferon formation using these translation systems? It is well established that the interferon mRNA is found in the polyadenylated RNA fractions, although there has been a report of activity also being associated with non-polyadenylated material (Montagnier *et al. Biochem. biophys. Res. Commun.* 59, 1031; 1974). No translatable RNA has been found in uninduced cells, and therefore induction is accompanied by a rapid increase in the amounts of translatable RNA. Thus the switch-on of interferon synthesis requires new transcription; the alternative explanation that translatable RNA is processed from a pre-existing RNA species is ruled out because of the well known inhibitory effect of actinomycin on interferon production. Two groups (Sehgal *et al. op. cit.*; Greene *et al. op. cit.*) have also shown that interferon mRNA is undetectable 6–8 h after induction with double-stranded RNA. This suggests that the switch-off of interferon production is due to cessation of mRNA

Translation of interferon messenger RNA *in vivo* and *in vitro*

from Derek Burke and John Morser

INTERFERONS are induced when a large variety of animal cells are treated with viruses or double-stranded polyribonucleotides. Cells which, before induction, were producing no detectable interferon, will produce up to 10^5 reference research units of interferon per 10^6 cells in the 24 hours after induction. Cells that have been induced are switched-off a few hours later. A second dose of interferon inducer does not lead to a second yield of interferon; the cells are described as hyporesponsive. Production of interferon can also be modulated in other ways; pretreatment of cells with small amounts of the homologous interferon before addition of an interferon inducer often increases the yield, a phenomenon termed 'priming'. Certain metabolic inhibitors added after induction with a double-stranded RNA can have a similar effect, a phenomenon called super-induction. Since purified interferon has a high specific activity (approaching 10^6 units per mg protein) and since the bioassay can easily detect 10 units of interferon, the production of very small amounts of interferon can be measured. Interferons are usually species-specific: interferon made in

mouse cells is inactive on human cells and *vice versa*. But even within the interferons produced by one species, there may be heterogeneity, for there are at least three different types of human interferon, which are distinguished by the type of cell in which they are formed: fibroblast interferon, leukocyte interferon and immune (type II) interferon produced when lymphocytes are treated with mitogens. Lymphoblastoid cells produce mainly the leukocyte type. These three interferons are antigenically distinct and have different physical and biological properties. The ease of detecting small amounts of interferon, the tight control of interferon production, and the existence of several different human interferons all make the interferon production system attractive for exploring the control of inducible proteins in eukaryotes. With the development of systems for translating interferon mRNA *in vitro* interesting answers about the control are beginning to emerge.

Three techniques have been developed for translating an interferon mRNA *in vitro*. The first involves application of RNA extracted from interferon-producing cells to heterologous cells with production of interferon characteristic of the cells from which the RNA was extracted (De

synthesis as well as to its inactivation.

The levels of interferon mRNA have also been measured in normal and superinduced cells, in order to determine whether the increased interferon production is due to a corresponding increase in the amount of mRNA. Raj and Pitha (*op. cit.*) found that the levels of interferon mRNA at 6 h after induction were very much higher in superinduced cells than in normal cells when the wheat germ system was used, but that there was little difference when the *Xenopus* oocyte system was used. In contrast Sehgal *et al.* (*op. cit.*) found the levels of mRNA were higher in superinduced cells using the oocyte system but they could not obtain translation of the same mRNA preparations in cell-free systems. Although there are differences between these two reports, both groups agree that the increased level of interferon production in superinduction is not due to the synthesis of new interferon mRNA, but is due to a change in the fate of the mRNA. Pitha and

Raj (*op. cit.*) suggest that similar levels of interferon mRNA are produced in normally induced and superinduced cells, but that the majority of the mRNA is in an inactive form in normally induced cells. They further suggest that superinduction is due to a modification of this RNA, which can also take place in *Xenopus* oocytes, but not in the wheat germ translation system.

However, Sehgal *et al.* (*op. cit.*) suggest that in normal cells mRNA is being degraded or inactivated but that superinduction prevents this process, through the action of a repressor gene. Although there are clearly still some difficulties in correlating the *in vivo* and the *in vitro* results, the experiments have already led to models of the control of interferon synthesis which can be tested. In the longer term, more detailed analysis will require other approaches, for example the use of nucleic acid hybridisation techniques to measure the amount of interferon mRNA directly. □

At that time, a number of endemic African forms moved out to occupy the Eurasian land mass. Hominoids of the *Proconsul* type were apparently among these emigrants and they rapidly spread very widely.

Recent work in Hungary (Kretzoi *Nature* 257, 578; 1975), Greece (de Bonis *et al. C.r. Séanc. hebdom. Acad. Sci. Paris Ser. D.* 281, 379; 1975; 284, 1393; 1977) Spain (Crusafont-Pairo & Golpe-Posse, *J. hum. Evol.* 2, 17; 1973), Turkey (Tekkaya *Bull. Min. Res. Exp. Ankara* 83, 148; 1974; Andrews & Tobien, *Nature* 268, 699; 1977) and on the Potwar Plateau has greatly increased our knowledge of this hominoid radiation.

It is now apparent that at least three genera, *Sivapithecus*, *Ramapithecus* and *Gigantopithecus*, may be recognised in deposits reflecting the more open woodland habitats of that time. These three genera share a complex of characters which may reflect a dietary adaptation to the harder or coarser foods of such environments. These features include thick enamel on the teeth, cheek teeth which were large in relation to body size and a mandible which was heavier and thicker than that of their more ape-like contemporaries.

From our point of view, the most interesting of these woodland-adapted genera is *Ramapithecus*. This genus is considered by many to be the earliest known member of the Hominidae, the family to which man and his phylogenetic ancestors belong. Indeed, as Pilbeam points out, unworn cheek teeth of this middle Miocene form very closely resemble those of later hominids such as *Australopithecus* and *Homo*. The earliest known member of this genus, and the family Hominidae, was recovered at Fort Ternan, Kenya just below a level dated with potassium/argon to 14 Myr BP. It has recently been suggested that an even earlier member of the genus may occur at Pasalar, Turkey (Andrews & Tobien *op. cit.*). However, the date of these Turkish deposits is based on faunal remains and it must be pointed out that not all elements of the Pasalar fauna confirm the suggested date of 15 Myr. For example, the bovid taxon *Prostrepsiceros rotundicornis* may occur at Pasalar but it does not appear elsewhere in Europe until about 10 Myr ago. If this species does indeed occur at Pasalar then a date for these deposits of 15 Myr must be treated with extreme caution.

However, the suggestion by Andrews and Tobien of a non-African origin for the hominids is an intriguing one. The genus *Ramapithecus* is known only from a single site in the whole of Africa and it is there found in association with a fauna that has many im-

Hominoid habitat shifts in the Miocene

from G. E. Kennedy

THE middle Miocene was a time of widespread environmental shifts which resulted in profound changes in many faunal groups. The tropical and subtropical forests of the earlier Cenozoic had provided large and virtually unbroken tracts of lush, protective habitat for many animals. But by the middle Miocene, around 9–14 Myr ago, a general cooling trend, coupled with tectonic and mountain building activity, had fragmented these forests into smaller, discontinuous habitat areas. With the decline of this habitat space, many animals moved from their long-occupied forest niches into the newly expanding, more open woodland areas. Among the Miocene faunal groups which made a shift from forest to open woodland were some groups of the Hominoidea—that primate super-family which contains apes and man, both living and fossil. In a recent issue of *Nature* (270, 684, 689; 1977) Pilbeam *et al.* report on some middle Miocene hominoids from the Potwar Plateau, in the Punjab Province of Pakistan which had made such a shift.

Hominoids of Miocene age have been known to palaeontologists since 1856. Their diversity, both morphological and geographical, led to a pro-

liferation of taxa which made the group unwieldy, if not incomprehensible, even to specialists. A revision of their taxonomy in 1965 (Simons & Pilbeam, *Folia primat.* 3, 81) was a welcome move in increasing understanding of the interrelationships of this group, which may contain the ancestors of living gorillas, chimps and man.

However, Pilbeam's work and the work of other primate palaeontologists now indicates that the elegant simplicity of that taxonomic scheme may have underestimated the vigour of the adaptive radiation of the middle Miocene hominoids. Since that revision several new genera have been proposed and more recently it has been suggested that two genera, previously sunk to sub-generic level, should now be reinstated as full genera. One of these, *Proconsul*, was an African early Miocene forest-living form, apparently at or near the ancestry of most later hominoids. Pilbeam now suggests that the other genus, *Sivapithecus*, was among the hominoids which made a habitat shift during the middle Miocene. A similar suggestion has recently been made by Andrews in a paper given at the VI International Primatological Congress, at Cambridge, 1976.

The opportunity to make this shift was enhanced when the long geographical isolation of the African continent ended about 16–17 Myr ago.

migrant, Eurasian elements. The geographical origins of the Hominidae should be a fascinating area of research in the coming years.

The environment in which the early hominids lived is also an intriguing area of study. Many theoretical discussions of hominid origins have been based on the development of an adaptive complex which emerged in response to pressures from savannah or grassland environments. However, as considerable recent research has indicated, grassland environments were not occupied by the earliest hominids and were, in fact, only entered by this group relatively recently. Therefore, occupation of the savannahs cannot have been the 'initial kick' which led to the emergence of the family Hominidae. Savannah adaptations may well have been important in the course of later hominid evolution, particularly in the Plio-Pleistocene radiation, but they cannot have been an initiating factor. Moreover, it is of interest and of uncertain significance that not all ramapithecines have been found in woodland deposits. In two areas, at Rudabanya, Hungary (Kretzoi *Nature op. cit.*), and, at Keiyuan, south China (Chow *J. paleont. Soc. Ind.* 3, 123; 1958) this form has been found in lignite deposits indicative of very moist, heavily vegetated conditions.

Clearly, the adaptive radiation of the middle Miocene hominoids was very complex. The human ancestors, rather than showing a unique and singular pattern of evolution were, in fact, part of a much broader process.

'Cardiac' RNA

from Jonathan Slack

It has long been realised by developmental biologists that the formation of ordered patterns of cell types during embryonic development requires the existence of interactions between different parts of the early embryo. The chemical nature of the putative signals has been a tantalising mystery ever since the abortive 'gold rush' to isolate Spemann's organiser during the 1930s. To the biochemically minded, informational macromolecules such as RNA present themselves as obvious candidates, while to the adherents of 'gradient' models they seem to be a red herring, and the signals are generally thought to be small molecules or ions which unleash complex responses

from the target cells.

Two recent papers from A. K. Deshpande and M. A. Q. Siddiqui (*Devl Biol.* 58, 230; 1977; *J. biol. Chem.* 252, 6521; 1977) make up an interesting contribution to this debate. They have isolated a specific fraction of RNA which causes a specific type of cell differentiation in explants from early chick embryos. At the stage concerned (Hamburger and Hamilton's stage 4), a chick embryo consists of two layers of cells: the epiblast above and the hypoblast below. Joining these layers down the centre is a mass of cells called the primitive streak through which cells are migrating to form a mesodermal layer in between the other two. At the front end of the primitive streak is a condensation of cells called Hensen's node, which is supposed to be the 'organising centre' for the formation of the general body plan.

Deshpande and Siddiqui have isolated their RNA species from the hearts of quite advanced (16 day) embryos. It is about 7S (200 nucleotides) in size and contains a sequence of poly(A). When the 'post nodal piece', roughly the posterior third of the early embryo, is treated with this RNA *in vitro*, it differentiates into cardiac muscle tissue.

What distinguishes this result from claims of this type in the past is the careful characterisation of the response and the large number of controls. The positive cases show beating movements and contain myofibrils (electron microscopy), actin and myosin (gels), glycogen granules (histochemistry) and acetylcholinesterase (assay) in quantities appropriate to cardiac muscle cells. The posterior explants do not form cardiac tissue or anything much else when cultivated on their own, nor when treated with RNA from a variety of other sources, or with synthetic polynucleotides. However, it is not yet proved that the 7S RNA enters the cells in an intact form.

The intriguing thing about this result is that it makes little sense in terms of early development. The fate of the post nodal piece in normal development is to form the posterior axial structures—notochord and somites. When cultured on its own *in vitro* it does not differentiate at all, and when cultured in the presence of Hensen's node it will differentiate into axial structures (Butros *J. exp. Zool.* 149, 1; 1962). It seems probable that this region is competent to form heart since the determination of the overall body plan is still going on at this early stage and classical work showed that quite an extensive lateral area would differentiate into heart when cultured *in ovo* on the chorioallantoic membrane (Rawles *J. exp. Zool.* 72, 271; 1936). But according to our present ideas about pattern formation the position of

the heart rudiment should be specified by two distinct signals. This is because the paired rudiments are situated in a two-dimensional sheet of cells and located symmetrically on either side of the axis. There is no particular reason why any factor involved in these early signals should be present in 16-d heart at all. Significantly enough, the authors show that brain, liver and kidney RNA are all inactive, both in stimulating heart differentiation and in stimulating differentiation of their own cell types.

The 7S RNA itself does not seem to be a mRNA even though it contains a poly(A) sequence, since it cannot be translated *in vitro* and actually inhibits the translation of purified messengers in cell free systems. It will be most interesting to see whether this species exists in the early as well as the late embryo, for if it has a role in normal development it should be possible to find it there. □

First interstellar NO bond

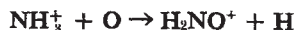
from Graham Richards

THE list of interstellar molecules grows longer and longer and many large polyatomic systems have been positively identified with still more postulated. Despite this, NO bonds have been conspicuously absent from the lists and NO itself has been the subject of many unsuccessful searches. Now with the detection of HNO by Ulrich, Hollis and Snyder (*Astrophys. J.* 217, L105; 1977) this anomaly in the chemistry of interstellar molecules may be removed.

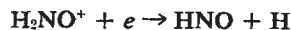
The observations were made in March, 1977 with the 36-foot radiotelescope of the National Radio Astronomy Observatory. The single line ascribed to HNO was detected in emission from the sources known to astronomers as SgrB2 and NGC2024. In general it is difficult to identify a new molecular species positively on the basis of a single spectral line but in this case very accurate laboratory frequencies are available (Saito & Takagi *J. molec. Spectrosc.* 47, 99; 1973). Further transitions of HNO may require observations at millimetre wavelengths.

Comparison with the very similar molecule HCO suggests an abundance ratio of carbon to nitrogen of about 1.4 in the same interstellar clouds, which conforms with cosmic abundance estimates. The chemical simi-

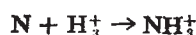
larities between HCO and HNO in all probability do not extend as far as the formation mechanisms. The formation of HCO is likely to depend on the availability of ionised carbon in moderate-density clouds. Ionised nitrogen may not be so readily available, so that the best suggestions for HNO formation come from Dalgarno



followed by



the initial NH_3^+ being derivable either from neutral nitrogen



or from the ion, N^+ , and unionised hydrogen molecules.

The letter in the *Astrophysical Journal* also reports a new unidentified line at a frequency of 81,505 MHz which may be a weak maser and observations of previously detected transitions of ketene (H_2CCO), dimethyl ether (CH_3OCH_3) and cyanoacetylene (HCCCN). A further unidentified line previously reported by Turner (*Astrophys. J. Lett.* **213**, L75; 1977) was also observed and ascribed to thioformaldehyde (H_2CS) or cyanoacetylene ($\text{HC}_2^{13}\text{CN}$).

The ambiguous identification, the unidentified line and above all the detection of an interstellar NO bond will all provide spurs for theoreticians to perform accurate calculations, spectroscopists to measure frequencies as well as encouragement for radio-astronomers to confirm and extend this study, perhaps once more searching for nitric oxide in interstellar regions.

Hydrothermal waters and heat flow

from Peter J. Smith

THE heat flowing outwards through the oceanic crust reaches the Earth's surface in three ways—by conduction, convection and radiation from the mantle and core, by conduction from radioactive sources in the crust itself, and by direct transport in the hot mantle material which rises at active oceanic ridges to form new lithosphere. Of the three, the most important is the last; indeed, more than 50% of the Earth's heat loss (that is, from oceanic and continental regions combined) takes place at oceanic ridges alone. With the possible exception of the few

Speed of light and relativity

from P. E. Hodgson

EINSTEIN'S special theory of relativity is based on two postulates—first that the laws of physics that hold in one coordinate frame also hold in any other coordinate frame moving with uniform velocity with respect to it, and second that the speed of light is independent of the velocity of the source. From these two postulates the theory of relativity can be derived.

The second of these two postulates can be tested experimentally and this has already been done in several ways. Among the most accurate is the study of gamma rays emitted from the decay of relativistic uncharged pions produced in very energetic nucleon-nucleon collisions. In one of these measurements the pions had an energy of 6 GeV and thus a velocity very nearly that of light, so that if the velocity of the source is added to that of the light, the gamma rays from their decay would have nearly twice the velocity of light, whereas if the velocities do not add the velocity is just c . There is thus nearly a factor of two between the velocities expected on the two possibilities, and this can be measured accurately by a time-of-flight method. If the result is expressed by saying that the measured velocity is

$$c' = c + kv$$

where k is a constant and v the velocity of the source, then the two possibilities correspond to $k = 0$ and $k = 1$. The measurements showed that k is smaller than 10^{-4} , which confirms the second postulate to high accuracy.

It is, however, always interesting to see if such an important result can be checked to even higher accuracy, and this has recently been done by Kenneth Brecher of the Center for Space Research at Massachusetts Institute of Technology (*Phys. Rev. Lett.* **39**, 1051; 1977). He made use of the regularly pulsing X-ray sources in binary star systems. Since in a binary star one component is orbiting round the other, if the system is observed in the plane of the orbit, for part of the period one star appears to be moving towards us and the other away from us, and the opposite happens half a

period later. If the velocity of the source is added to that of the light then we would not see two stars rotating around each other but instead would see complicated effects depending on the distance of the stars. In particular we would not see them continuously; they would appear and disappear as the various beams from the stars in different positions overtake one another as a result of their different velocities.

This method of confirming that the velocity of light is independent of that of the source was suggested long ago, but for light in the visible region it suffers from the objection that the radiation from a distant star is scattered as it goes through the intervening interstellar matter, and is then replaced by reradiated fields produced by the dipoles of the medium. This wave is then propagated with the phase velocity of the medium. Calculations show that for light in the visible region the characteristic length for the extinction of the original radiation is about two light years, which is less than the distance of the nearest star.

This objection does not apply to the 70 keV X rays emitted from some binary star systems, as they are much more penetrating and the interstellar extinction is negligible. Such stars thus provide a very sensitive way of testing the second postulate.

Brecher applied this method to three binary X-ray sources, and determined their distances, orbital periods, and Doppler velocities. A small but finite value of the constant k would affect the eclipse times and the apparent eccentricity of the orbit, and so measurements of these quantities enable upper limits to be set to k . A careful analysis of the results of the measurements showed that for one of the stars k is less than 2×10^{-9} , and for the other that it is less than 4×10^{-10} . These limits are much less than those set by the pion method. The second postulate is thus established to much greater accuracy than ever before.

P. E. Hodgson is a lecturer in Nuclear Physics in the University of Oxford.

local continental areas of high heat flow capable of providing geothermal power, oceanic ridges are therefore the most interesting of all the Earth's thermal features. They are also the most important in the sense that study of their thermal characteristics is capable in principle of throwing light on

some of the details of the plate tectonic processes which brought the heat to the surface in the first place.

Unfortunately, however, oceanic ridges are precisely those regions where heat flow measurement is most difficult and the results most uncertain. Conventional heat flow probes require for

their operation at least several metres of sediment, but the new lithosphere in ridge zones has had little time to develop a sediment cover. Measurement is therefore limited to the small isolated basins where sediment has begun to form, which are by definition not typical of the whole region. Moreover, the rugged ridge topography, which allows such sediment pools to grow and which is itself the product of variable thermal expansion, distorts the flow of heat and thus introduces geographical variations that are difficult to correct for. Small wonder, therefore, that heat flow data from ridges, though generally high as predicted by plate tectonics, are both highly scattered and subject to systematic errors which generally do not even permit calculation of a true mean value.

Nor is that all. Some of the systematic deviations (for example, the conspicuously low heat flow often found on ridge flanks) are so great that it is impossible to account for them either by topographic distortion or as the result of the accumulation of cold sediment. This has led several workers to suggest that an important influence on the distribution and variation of heat flow in the vicinity of oceanic ridges is hydrothermal circulation in the igneous upper crust—an idea put forward many years ago by Elder (in *Terrestrial Heat Flow*, edit. by Lee, W. H. K., American Geophysical Union, 1965) but promoted particularly vigorously in recent years by Lister (*Geophys. J.* **26**, 515; 1972 and subsequent articles). Lister's point is that heat advected from the crust by hydrothermal circulation which has open access to the ocean water will not be recorded by measurements limited to sediment pools. Average heat flow values obtained from ridge zones will therefore almost always be too low to be consistent with otherwise reasonable models of lithospheric accretion which ignore hydrothermics. Only when there is a complete sediment cover forming a continuous impermeable barrier between the igneous crust and the open ocean will the measured average heat flow begin to approach the true value of heat lost by the oceanic lithosphere.

The existence of hydrothermal circulation is now widely accepted, having been demonstrated to be the most likely explanation for the heat flow distributions revealed by several studies since 1972. A particularly striking confirmation of Lister's basic ideas, however, has now been reported by Davis and Lister (*J. geophys. Res.* **82**, 4845; 1977) who have obtained 104 heat flow results from a 110 km square area centred on the point where the Juan de

Fuca ridge meets the transform (Sovanco fracture zone) that offsets it from the Explorer ridge to the north. In the southern section of this area lies the northern end of the Juan de Fuca ridge with its young lithosphere (0–1.8 Myr old), whereas to the north of the fracture zone lies the older crust (3.3–3.6 Myr) which has spread from the Explorer ridge. In other words, the test area includes two quite distinct zones between which the effects of hydrothermal circulation, if any, may be compared.

Throughout the area as a whole the heat flow observations, made at roughly 10 km intervals, were highly variable, ranging from 1.4 to 15.9 $\mu\text{cal cm}^{-2} \text{s}^{-1}$ and having no obvious correlation with either topography or sediment cover. Moreover, this variability and lack of correlation were equally evident in eight measurements made specially at 2 km intervals. From this evidence alone it seems necessary to invoke hydrothermal circulation throughout the test area.

But the consequences of the circulation differ significantly between the two subzones. The overall picture is mirrored south of the fracture zone where the sediment distribution over the Juan de Fuca ridge and ridge flanks is intermittent and the sediment, where it exists at all, is thin. The heat flow variability is far too great to be explained solely in terms of topographic distortion or sedimentary accumulation. Where there is no sediment cover there must be open circulation between the hydrothermal waters in the igneous lithosphere and the ocean waters above, leading to an underestimate of the true heat flow by the actual measurements made in sedimentary areas. This is confirmed by plotting the measured mean heat flow against the reciprocal of the square root of the age of the lithosphere. The point thus produced lies way below the line derived from other Pacific heat flow provinces where there is no hydrothermal circulation.

North of the Sovanco fracture zone, on the other hand, the basement of the east flank of the Explorer ridge is completely covered with thick flat-lying sediment. The heat flow here is still variable, but rather less so than in the south of the test area; and there is now a distinct inverse correlation between heat flow and the corresponding sediment thickness. In principle, both the variability and the correlation could be explained either by topographic distortion (by which heat tends to converge towards areas where, because of topographic variations in the basement, the sediment is thinner and hence its insulating effect smaller) or by the cooling influence of sedimentary deposition. Indeed, the presence of the heat flow-sediment thickness correlation in par-

ticular points strongly towards just such explanations. Calculations show, however, that even in this zone these effects are insufficient.

So it is necessary to invoke hydrothermal circulation even though in this case there is a complete sediment cover. But this is not to say that the sediment cover has no effect. On the contrary, it prevents the hydrothermal waters in the igneous lithosphere venting to the ocean and thus eliminates the false measurement of heat flow inevitable in the case of the Juan de Fuca ridge zone. The measured average heat flow is thus the true heat flow; and this is confirmed by the fact that the former plots precisely on the Pacific heat flow-lithospheric age curve.

In summary, then, what Davis and Lister have shown, above all, is that artificially low values of measured heat flow are entirely consistent with the effects of hydrothermal circulation but that such circulation does not give rise to artificially low values of measured heat flow in all circumstances. Although hydrothermal circulation will always produce scatter of heat flow, a false average value will result only when the hydrothermal waters can vent to the ocean. Incidentally, such venting will, not surprisingly, also lead to a reduction in the average temperature of the hydrothermal waters themselves. Thus Davis and Lister estimate that temperatures at the top of the igneous crust range from 300 °C to 100 °C below the few well-sedimented sections of the Juan de Fuca ridge but from 200 °C to below 25 °C where there is no sediment cover. □



A hundred years ago

THE vessel was at this time in the neighbourhood of Terra del Fuego, about 140 miles from Magellan's Straits, when early in the morning it narrowly escaped collision with an island where no trace of land appeared on the charts. The vessel hove-to until daylight, when the captain proceeded with a boat's crew to the new island, which had gradually diminished in size since the first observation. Around the conical rocky mass the water was hissing, and although no smoke appeared, it was found to be too highly heated to permit of landing. The sinking continued slowly, until at eight o'clock the island was completely submerged, and an hour later the vessel passed over the spot where it had disappeared.

From *Nature* 17, 3 January, 194; 1878.

review article

Prediction of protein structure from amino acid sequence

Michael J. E. Sternberg* & Janet M. Thornton*

In principle, it is possible to predict theoretically the three-dimensional structure of a protein from its amino acid sequence. Recently substantial progress towards this goal has been made by the use of simple models to represent protein conformation and interatomic interactions, together with the knowledge gained from analyses of known protein structures.

THE distinguishing feature of biological macromolecules, especially proteins, is their ability to adopt a unique three-dimensional structure (Fig. 1). It is this precise structure which is essential for recognition between different molecules in an organism and determines how they function. Thus knowledge of tertiary structure is vital for understanding biological function. Experimental evidence shows that all the information necessary for a protein to acquire its complex structure is stored in its linear amino acid sequence¹. Therefore theoretically one should be able to predict the conformation of a protein from a knowledge of sequence alone. Once a protein has been isolated, its sequence can almost always be determined. In contrast, structure determination is difficult and many proteins defy all attempts to obtain suitable crystals. Thus a successful prediction of three-dimensional structure, which implies a basic understanding of how structures are formed, would be an important step forward in molecular biology.

This review will examine the progress made in the prediction of protein structure from amino acid sequence over the last five years (for other reviews see refs 2–4). The major problem of prediction is that even a small protein, 50 amino acids, could adopt about 10^{50} conformations. How does one find the native structure amongst the myriad of alternatives? Recent progress towards a solution of this problem has been made by use of simple models to represent both the conformation and interatomic interactions of the protein, together with the knowledge gained from the increasing number of protein structures known.

Thermodynamics and kinetics of refolding

Many experiments^{1,5} have shown that most proteins can spontaneously refold in a biological environment, from a denatured state to their native conformation. Generally this is true for proteins with and without disulphide bridges and for both monomeric and multi-subunit proteins. These experiments led to the thermodynamic hypothesis¹, which states that the polypeptide chain folds so as to attain the conformation of lowest free energy. This hypothesis was modified in 1968 by Levinthal⁶, who also considered the kinetics of folding. *In vivo* proteins require between 10^{-1} and 10^3 seconds to fold up, and, since molecular rearrangements occur in about 10^{-13} s, a protein cannot possibly search all conformations ($> 10^{50}$) (ref. 7). Thus folding must be directed along pathways and the native structure will be at the minimum free energy of the kinetically accessible conformations. This local minimum need not necessarily correspond to the global minimum of free energy. To reduce the number of conformations searched it is thought that portions of the chain serve as

nucleation sites, around which the remainder of the protein folds. The α -helix and β -strand are possible sites. One suggestion for the pathway, called N-terminal folding, is that the polypeptide chain folds as it is being synthesised from its N-terminus⁸. Definitive experimental evidence on folding pathways has proved difficult to obtain. Studies on pancreatic trypsin inhibitor⁹ and hen egg white lysozyme¹⁰ suggest that at most there are only a limited number of folding pathways to the tertiary structure.

Energy calculations for atom–atom interactions

Theoretical predictions of structure have been influenced by both the thermodynamic and kinetic arguments. Initially the thermodynamic properties of polypeptide chains were considered. Liquori¹¹, Ramachandran¹² and their coworkers showed that the peptide unit can adopt only certain allowed conformations (Fig. 1d). By considering non-bonded interactions, they constructed ϕ – ψ plots which have subsequently been improved by semi-empirical energy calculations¹³ to predict preferred regions. There is good agreement between these preferred conformations and the observed distribution of ϕ , ψ angles. Recently energy calculations have been refined, either by choosing parameters to fit crystal structures¹⁴ or by performing complex molecular orbital calculations¹⁵.

It follows from the thermodynamic hypothesis, that by calculating the total free energy of a protein and finding the global minimum, it should be possible to predict the native structure. But at present this approach is not feasible. Not only is the computer time required excessive, but the energy surface is extremely complex with numerous minima. Consequently studies have been restricted to small sections of proteins and polypeptides¹⁶, and the refined calculations have only been valuable in the understanding of local conformation. The multi-minimum problem will probably be solved by finding the correct folding pathway. One step towards this goal would be the prediction of regular secondary structures from the sequence, since these structures may form the nucleation sites (see for example ref. 17).

Prediction of secondary structure

More than fifteen methods of predicting secondary structure from amino acid sequence have been proposed. (References are given to some of the more recent approaches: 18–27.) These predictions assume that local sequence determines local structure, which has been partially verified by the results. Usually only short-range interactions are considered, interactions between distant chain segments being ignored.

Broadly the methods can be divided into two categories. First, there are the statistical methods. In proteins of known sequence and structure, each residue is assigned to one class of secondary

*Laboratory of Molecular Biophysics, Department of Zoology, South Parks Road, Oxford, UK

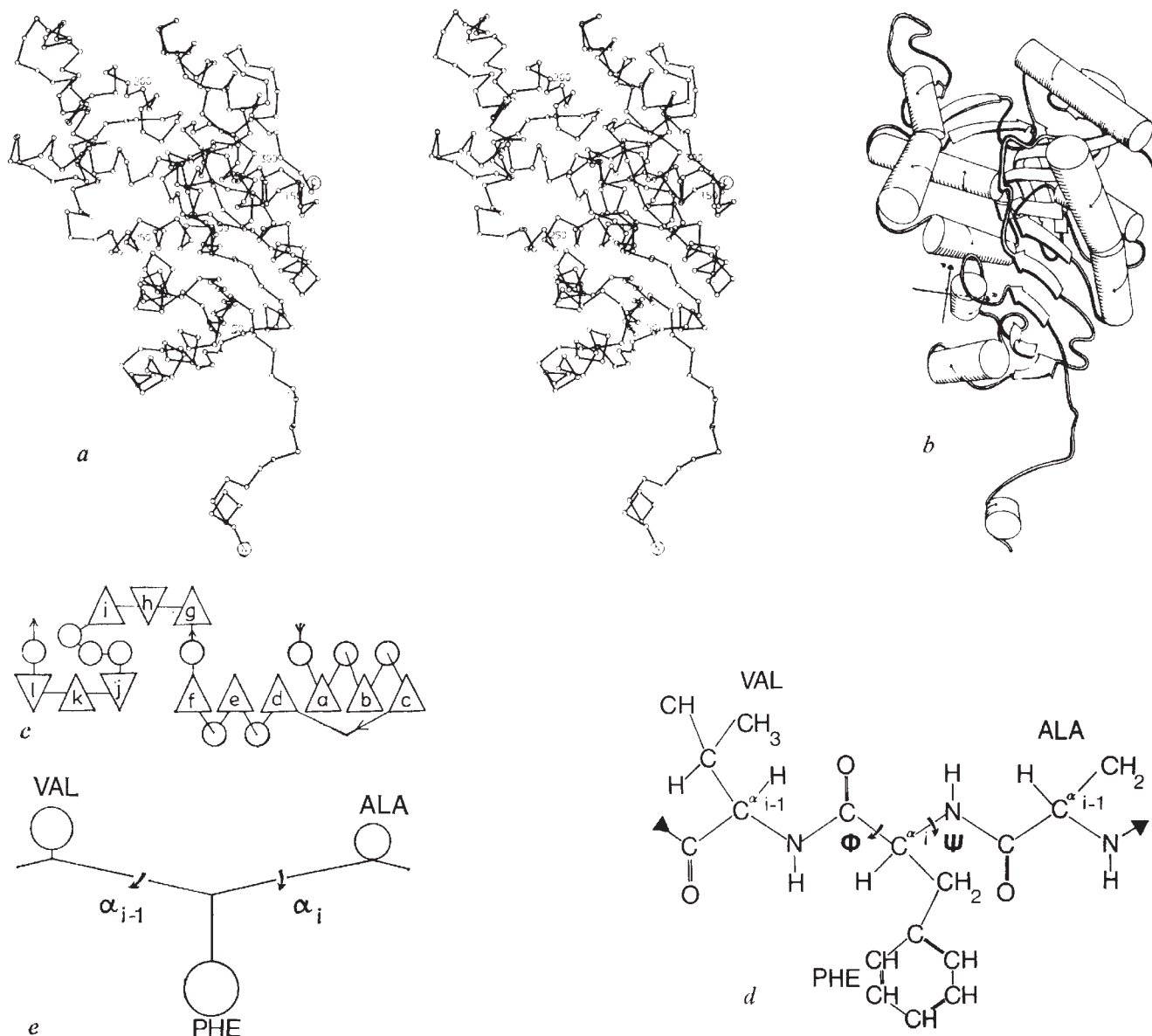


Fig. 1 Protein Structure. The figure shows how the complex three-dimensional structure of a protein—(lactate dehydrogenase (LDH))—can be simplified for use in analysis and prediction. *a*, A stereo diagram of the main chain (C² atoms) of LDH (to be viewed with stereo glasses). *b*, Schematic diagram of the main chain of LDH in the same orientation as Fig. 1*a*, showing the relative positions of the α-helices (cylinders) and β-strands (ribbons). *c*, A highly schematic diagram of LDH showing the relative positions of α-helices (○) and β-strands (Δ and ▽ according to whether viewed from the N- or the C-terminus of the strand). *d*, A section of polypeptide chain showing all the atoms and the dihedral angles φ and ψ. *e*, Simplification of (*d*) used by Levitt and Warshel^{52,55} to predict structure.

structure: α-helix, β-strand, coil, and sometimes β-turn. The observed distribution of amino acids amongst these classes is then used to derive parameters for the probability that a particular amino acid type will adopt a given class of secondary structure. The methods vary both in how the parameters are derived and how they are then used to predict structure. Some schemes sum the effects of individual residues^{18–20} while others also incorporate pair and/or triplet information^{21–24}. The method of Chou and Fasman¹⁹ has attracted most attention as it is simple to understand, can be applied without a computer, and has been relatively successful.

Second, there are a few methods^{25–27} which do not incorporate statistical data but are based on stereochemical criteria. Lim²⁷ has considered both the hydrophobicity and size of side chains to propose favourable patterns of residues that will form α-helices and β-strands. This apparently complex method, once understood, can be applied easily without a computer. It is appealing because it emphasises the importance of the positions of hydrophobic and hydrophilic residues and also incorporates, to some extent, long-range interactions.

Schulz²⁸ and Matthews²⁹ combined the results of several predictions to obtain joint prediction histograms for adenyl kinase and bacteriophage T4 lysozyme, respectively. The joint predictions were shown to be comparable with the best single prediction. This combined approach has been computerised by Argos *et al.*³⁰ who then applied it to 20 independent protein structures. An advantage of this completely computerised method is that it is impartial, whereas some of the individual schemes require human intervention to make the final assignments. The following generalisations about secondary structure prediction can be made.

No individual method is clearly superior. The success of a prediction for helices can be measured by a correlation coefficient C_α (ref. 29), where $C_\alpha = 1$, indicates a perfect prediction, while $C_\alpha = 0$, indicates a random prediction. C_β and C_T are similarly defined for predictions of β-strands and β-turns, respectively. The success of a given scheme varies considerably between proteins. Better predictions are generally obtained for small thermostable monomeric proteins than for the larger molecular weight structures. Argos *et al.*³⁰ find a range of values of $C_\alpha = 0.2$ to 0.86,

$C_{\beta} = 0.03$ to 0.73 and $C_T = 0.10$ to 0.84 . Thus the predictions were better than random for all structures. Although most of the α -helices and more than half of the β -strands and β -turns are located roughly, there are often errors in determining their termini. Presumably α -helices are predicted more accurately as local interactions are particularly important in their formation. Often the amino-terminal half of protein is predicted better than the carboxyl-terminal half, suggesting an amino-terminal nucleating core, in which short-range interactions dominate.

The success of predictive schemes has not greatly improved since 1974 although more protein structures are now available. This suggests that failure to achieve 100% accuracy is not the result of limited data, but that the formation of secondary structure is also determined by long-range interactions, which are not adequately considered.

Analysis of tertiary structure

The globular tertiary structure of a protein is stabilised by long-range interactions, which occur between non-local sections of sequence, for example the packing between two helices. With the increase in the number of protein structures known, many common structural features have been recognised, which were not previously predicted, and it has become apparent that proteins follow certain 'folding rules', which help determine tertiary structure. The simplification of the complex tertiary structure by the use of schematic diagrams (Fig. 1*b, c*) has helped in the understanding and comparison of protein structures. The results of analyses of tertiary structure, which are important for structure prediction, will be described.

Inspection of detailed interatomic interactions shows that two features are important determinants of protein structure. All proteins are closely packed, with packing densities comparable to those of small organic crystals³¹. This requirement for close-packing restricts the observed packing angles of α -helices and β -sheets to a few categories³². Second, hydrophobic groups tend

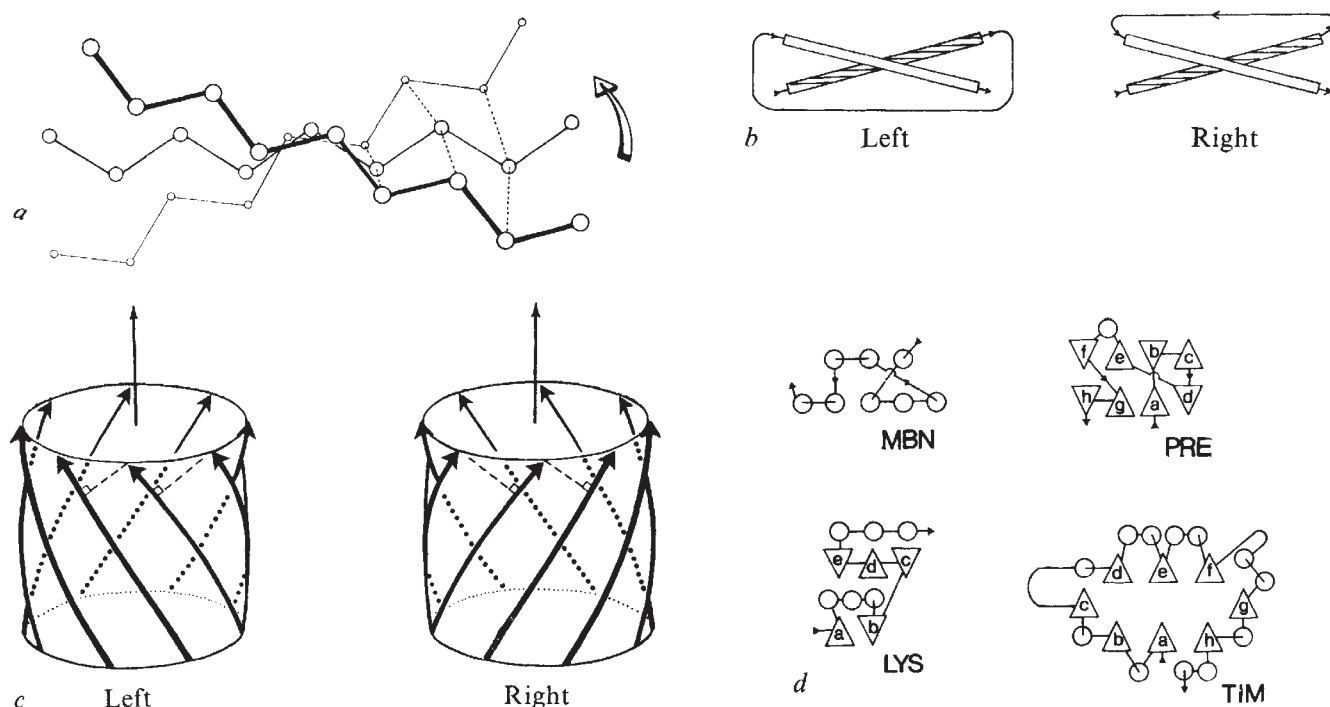
to be buried inside the protein forming a non-polar core, while hydrophilic side chains are nearly always exposed³³. Nearly all the buried polar groups, such as the main-chain $-\text{NH}$ and $-\text{CO}$ groups form hydrogen bonds³⁴. Consequently chain reversals almost always occur near the surface of the molecule³⁵ while large parts of the α -helices and β -sheets are buried. Any realistic prediction must aim to satisfy both these requirements.

Further restrictions on possible conformations occur because of the chirality of L-amino acids. In particular: (1) α -helices are right-handed. (2) β -sheets have a right-handed twist of the peptide units along the β -strand and, consequently, a left-handed rotation between adjacent strands³⁶ (Fig. 2*a*). (3) The connection between two parallel strands in the same sheet will almost certainly be right-handed (Fig. 2*b*)³⁷⁻³⁹. (4) Cylindrical sheets (that is β -barrels) are right-handed (Fig. 2*c*)⁴⁰. The first requirement for right-handed α -helices can be incorporated into a conventional energy minimisation (see later), but the remaining observations refer more specifically to tertiary structure and suggest that a model-building approach to protein prediction could be useful.

One of the main difficulties in predicting tertiary structure is to determine which segments of chain will lie close together. It has been observed that elements of secondary structure which are sequential along the polypeptide chain tend to lie proximate in the three-dimensional structure⁴¹. Furthermore analyses of β -sheets have suggested several preferences which help determine strand order^{42,43}. For example, the last sequential strand in the sheet preferentially occupies an edge position. Similarly the most hydrophobic strands generally occupy central positions in the sheet, with the more hydrophilic strands at the edge⁴⁴. These preferences should also be incorporated into prediction schemes (see later).

Several features of the overall fold have been recognised. Certain tertiary structures of more than an individual α -helix or β -sheet are often observed and appear to act as 'building blocks' or 'super-secondary structures'⁴⁵. Large proteins are often divided into a few distinct structural regions, called domains. These

Fig. 2 Results of analysis of protein structure. *a*, A twisted β -sheet: schematic diagram of three β -strands viewed normal to their direction to show the left-handed rotation between them that is always observed³⁶. *b*, Two parallel β -strands, viewed normal to their direction, and a connecting region. The right-handed structure is nearly always observed whereas the left-handed structure occurs rarely³⁷⁻³⁹. This may be because the left-handed rotation between β -strands gives a direct connection for the right-handed structure³⁷. *c*, The handedness of the β -barrel. In all observed barrels the strands are inclined to the axis so as to form a right-handed structure⁴⁰, that is if the strands were extended they would tend to form a right-handed helix about the barrel axis. This handedness is a consequence of the twist of the β -sheet. The alternative left-handed structure has not been observed. *d*, Classes of protein structure⁴¹. MBN-myoglobin, an α -protein; PRE-prealbumin, a β -protein; LYS-hen egg white lysozyme, an $\alpha + \beta$ protein; TIM-triose phosphate isomerase, an α/β protein.



domains contain 40–250 residues. Finally most protein structures fall into five rather distinct classes⁴¹ which are defined according to secondary structure content (Fig. 2d): α -proteins have only α -helix; β proteins have mainly β -sheet; $\alpha + \beta$ proteins have α -helix and β -strand secondary structural segments that do not mix but tend to segregate along the polypeptide chain; α/β have mixed or approximately alternating segments of α -helix and β -strand; and 'coil' proteins contain almost no regular secondary structure. We have noted that intracellular proteins usually fall into the α or α/β classes, and almost never have disulphide bridges. In contrast, extracellular proteins are usually all β or $\alpha + \beta$ type.

Once a protein sequence is known and a secondary structure prediction has been performed, these observations will help to predict the correct tertiary structure. Although some of these results have already been used, most of this work has yet to make its impact on prediction.

Tertiary structure prediction from evolutionary relationships

Families of homologous proteins, for example the cytochromes c, have almost identical main-chain structures but only similar sequences, suggesting that structure evolves more slowly than sequence⁴⁶. Observed structures can therefore be used to predict unknown structures of proteins with related sequences. For example, the fold of α -lactalbumin has been correctly predicted on the basis of its sequence similarity with hen egg white lysozyme⁴⁷. Similarly, a structure for troponin C was predicted from the crystal structure of myogen in conjunction with symmetry requirements⁴⁸. Recently, sequences that show very little similarity have been observed to adopt similar fold, for example in the case of the dehydrogenases⁴⁹. This has proved useful for prediction, although it is debatable whether these similarities result from convergent or divergent evolution. For example, parts of aldolase⁵⁰ and glutamate dehydrogenase⁵¹ have been predicted to fold in a similar six-stranded structure to lactate dehydrogenase. This is based on secondary structure prediction, sequence resemblances with the dehydrogenases and, for aldolase, recognition of the fold by affinity chromatography⁵⁰.

One can envisage that grouping proteins into families may be used widely to predict structure. Furthermore many proteins have evolved by gene replication or gene fusion, and the structural consequences of this may also be useful for prediction. However, this type of approach is necessarily restricted to related proteins.

Prediction of tertiary structure

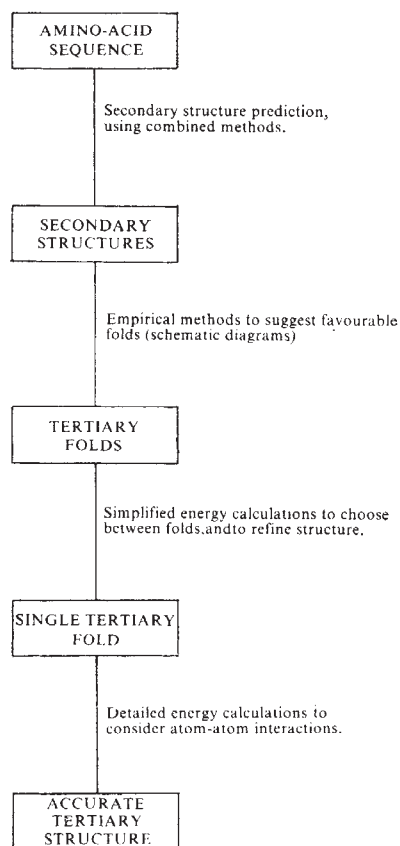
Predictions of the tertiary structure of a protein by more general methods have been made only very recently. At this early stage the predictions have been limited to small proteins whose structures are already known. There are several completely different approaches.

The semi-empirical energy calculations described previously have been modified to treat proteins by a simplification of energy functions and protein structure representation (Fig. 1e), followed by energy minimisation^{52–55}. These methods are based on the assumptions that: (1) the fine details of protein structure can be time-averaged and are not vital for the formation of the correct fold of the polypeptide chain; (2) only variables that have the greatest effect on folding need be considered; (3) the major determinants of the protein fold are the requirement for close-packing (van der Waals interactions), good hydrophobic burial and hydrophilic exposure (solvent interactions) and sometimes the ϕ - ψ preference of certain amino acid types. Probably the choice of these determinants was influenced by the analyses of known structures. These simplifications were introduced by Levitt and Warshel⁵². They represented each amino acid residue by two centres, the C α atom and a single sphere positioned at the centroid of the side chain (Fig. 1e). The interactions were between these centres. Only one degree of freedom—the torsion angle α about a line joining consecutive C α atoms—was allowed per residue⁵⁵. Minimisation of the energy function was followed by 'thermalisation' to escape from local minima. The method was

applied to pancreatic trypsin inhibitor (58 residues with one α -helix of ten residues) and myogen (108 residues with six α -helices). Generally residues that formed α -helices were preset to their correct conformation and held rigid during the calculation. A similar algorithm has been developed by Kuntz *et al.*⁵³ who minimised an error function which incorporated hydrophobic/hydrophilic interactions, steric repulsion and correct disulphide and iron-sulphur bonding. The structure of trypsin inhibitor and rubredoxin (53 residues) were predicted. Tanaka and Scheraga⁵⁴ have outlined a third method which shares some features with the two approaches described above. Each side chain is represented by a single sphere and the interactions between pairs of residues is estimated from an analysis of close contacts in 25 proteins of known structure. The secondary structure for each residue was predicted. Low energy conformations were found by a random 'Monte-Carlo' search, rather than by minimisation. The method was applied to trypsin inhibitor.

The success of these methods can be measured as the r.m.s. deviation of the predicted structure coordinates from the observed structure. Levitt and Warshel obtained r.m.s. errors of 5–7 Å for trypsin inhibitor. Conformations that deviated more from the native structure always had higher energies, which gives an independent criterion for choosing the best conformation. They had less success with myogen, obtaining only three structures with r.m.s. values less than 8.5 Å, even assuming rigid helices. Similarly Kuntz *et al.* obtained a number of structures for each protein with r.m.s. errors of 4 to 6.5 Å for trypsin inhibitor and rubredoxin (note: the disulphide and sulphur-iron bonding was forced to be correct). In both methods some, though by no means all, of the minimised structures had the correct fold, with bends and loops approximately located and the correct long-range interactions. Myogen was the most noticeable exception, since even the structures with lowest r.m.s. errors showed an incorrect fold. Tanaka and Scheraga do not give r.m.s. values but claim to have obtained some agreement with the native structure.

Fig. 3 A possible approach to predicting tertiary structure from amino acid sequence using a combination of methods.



Several serious points arise from consideration of these results. These methods have not yet been sufficiently tested for it to be clear whether they will be generally useful and whether the assumptions made are valid. Their application to larger proteins is difficult to envisage since the computing time will increase rapidly and larger structures are more complex. Levitt (personal communication) has suggested that compact structures of trypsin inhibitor, generated by a self-avoiding random walk, are almost as close to the native structure as some of the better structures obtained from energy minimisation. Thus it is questionable whether the simple structural representation and energy minimisation techniques are adequate. Finally, it is difficult to see how many of the empirical observations on tertiary structure can be incorporated into this kind of approach.

Overall, the results of these simplified energy calculations were very encouraging. Starting from a fully extended conformation, proteins have been folded into compact structures which show many similarities to the native states. It is to be hoped that these methods will be more fully explored and refined to give improved predictions.

A different type of approach to predict the structure of α -helical proteins has been described by Ptitsyn and Rashin⁵⁶, used by Lim and Efimov⁵⁷, and modified for application to β -sheet proteins by Zav'yalov⁵⁸. There are two basic assumptions. First, folding proceeds by a stepwise process in which fluctuating regions of secondary structure are formed and then condense into a single compact globule. Second, the condensation process is determined by three considerations: (1) Regions of secondary structure, which are adjacent along the polypeptide chain, will tend to lie proximate in the tertiary structure. (2) At each stage of the folding pathway only the most favourable structures are maintained and all other structures rejected. (3) Packing will occur so as to maximise the reduction of hydrophobic free energy, by the exposure of hydrophilic groups and burial of hydrophobic side chains.

Thus Ptitsyn and Rashin⁵⁶ took the residues forming α -helices in myoglobin from the native structure and modelled these α -helices by cylinders. A manual search was performed for the packing between adjacent helices which gave optimal reduction in hydrophobic free energy. They were able to predict the correct native fold. Similarly Lim and Efimov have folded tobacco mosaic virus protein. Zav'yalov⁵⁸, assuming the hairpin bends in the immunoglobulin domain, predicted the structure of the β -sheet sandwich. The success of this approach indicates the importance of hydrophobic free energy and close-packing of secondary structures in determining tertiary structure. The major problem is that the method is non-automated and therefore difficult to apply impartially. However a computerised method based on this approach, which incorporates the packing rules for secondary structures³² and includes some calculation of hydrophobic free energy could prove successful.

Most recently a different approach has been considered in which, rather than energy calculations, folding rules obtained from the analyses of structures are used for prediction. We have used the results from a quantitative analysis of β -sheets to derive a folding parameter which represents whether a particular arrangement of strands in a sheet is favourable⁵⁹. This approach involves a further level of simplification: the calculations only consider the relative positions and type of connections of the strands (as in Fig. 1c) and do not evaluate residue-residue interactions. To explore this approach, the location of the secondary structure along the chain and the assignment of strands to a particular sheet were taken from the observed structure. For 25 out of the 52 sheets examined, the actual arrangement is the sheet, or one of the sheets, with the most favourable folding parameter and therefore could have been readily predicted. This surprisingly good result suggests that strand order is somewhat determined by a few elementary folding rules.

Nagano (personal communication) has suggested another empirical approach to predicting probable packing arrangements of α -helices and β -strands in α/β type proteins. First, regions of the sequence that are likely to be predicted as α or β secondary

structures are located. Then hypothetical rules are used to suggest packing arrangements whose stability is evaluated by an empirical penalty function. Thus he builds up a schematic diagram of the protein structure (like that in Fig. 1c). The generality of this method has still to be determined. However it is a novel approach, which has the advantage of starting from the sequence and translating right through to the tertiary structure.

These empirical methods, by their simplicity, can be applied to larger proteins whose size puts them beyond the scope of conventional methods. The approaches consider the structural features which result from both the stability of the final structure and the kinetics of protein folding, without any assumptions about the folding process. They should be useful to predict probable folds which then can be examined by energy calculations.

Outlook

Although the individual methods of structure prediction will doubtless be extended and improved, it is likely that the answer to the folding problem lies in a combination of the available approaches. One combination is outlined in Fig. 3. This shows how, starting with the sequence, one can progress through the secondary structure analysis and the prediction of tertiary fold to simplified and then detailed energy calculations. Feedback at each stage is important, so that, for example, the final choice of secondary structure must be influenced by consideration of which tertiary folds are favourable. For an individual protein, biochemical data (for example, which residues are involved in the active site) can be incorporated, either to suggest which of the tertiary folds should be chosen, or to test the feasibility of the final predicted structure. Perhaps the major improvements to prediction will come from further renaturation and X-ray studies, which will provide more information on folding pathways and the final structure. Clearly the last few years have witnessed the development of many novel approaches to the prediction of tertiary structure. Hopefully these preliminary investigations will provide the first steps towards the solution of the folding problem.

We thank: Professor D. C. Phillips for encouragement and advice; Dr I. A. Wilson for help in preparation of the stereo diagram; Dr Olga Kennard for supplying coordinates of lactate dehydrogenase from the Cambridge Data Bank; Professor M. G. Rossmann and Academic Press, New York, for allowing us to reproduce Fig. 7 from reference 49; and the Medical Research Council for financial support.

1. Anfinsen, C. B. *Science* **181**, 223-230 (1973).
2. Scheraga, H. A. *Adv. phys. org. Chem.* **6**, 103-184 (1968).
3. Ramachandran, G. N. & Sasisekharan, V. *Adv. prot. Chem.* **23**, 283-437 (1968).
4. Némethy, G. & Scheraga, H. A. *Quart. Rev. Biophys.* (in the press).
5. Wetlaufer, D. B. & Ristow, S. *Ann. Rev. Biochem.* **42**, 135-158 (1973).
6. Levinthal, C. J. *Chim. phys.* **65**, 44-45 (1968).
7. Wetlaufer, D. B. *Proc. natn. Acad. Sci. U.S.A.* **70**, 697-701 (1973).
8. Phillips, D. C. *Proc. natn. Acad. Sci. U.S.A.* **57**, 484-495 (1967).
9. Creighton, T. E. *J. molec. Biol.* **113**, 275-293, 295-312 (1977).
10. Anderson, W. L. & Wetlaufer, D. B. *J. biol. Chem.* **251**, 3147-3153 (1976).
11. De Santis, P., Giglio, E., Liguori, A. M. & Ripamonti, A. *Nature* **206**, 456-458 (1965).
12. Ramachandran, G. N., Ramakrishnan, C. & Sasisekharan, V. *J. molec. Biol.* **7**, 95-99 (1963).
13. Lewis, P. N., Momany, F. A. & Scheraga, H. A. *Israel J. Chem.* **11**, 121-152 (1973).
14. Hagler, A. T., Huler, E. & Lifson, S. *J. Am. chem. Soc.* **96**, 5319-5327 (1974).
15. Pullman, B. & Pullman, A. *Adv. prot. Chem.* **28**, 347-526 (1974).
16. Isogai, Y., Némethy, G. & Scheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 414-418 (1977).
17. Robson, B. & Pain, R. H. *Biochem. J.* **155**, 331-344 (1976).
18. Ptitsyn, O. B. & Finkelstein, A. F. *Biofizika (USSR)* **15**, 757-767 (1970).
19. Chou, P. & Fasman, G. D. *Biochemistry* **13**, 212-222; 222-245 (1974); *ibid. Adv. Enzym.* (in the press).
20. Burgess, A. W., Ponnuswamy, P. K. & Scheraga, H. A. *Israel J. Chem.* **12**, 239-286 (1974).
21. Robson, B. & Suzuki, E. *J. molec. Biol.* **107**, 327-356 (1976).
22. Maxfield, F. R. & Scheraga, H. A. *Biochemistry* **15**, 5138-5153 (1976).
23. Nagano, K. *J. molec. Biol.* **109**, 251-274 (1977).
24. Kabai, E. A. & Wu, T. T. *Biopolymers* **12**, 751-774 (1973).
25. Kotelchuck, D. & Scheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **61**, 1163-1170 (1968).
26. Lewis, P. N., Go, N., Go, M., Kotelchuck, D. & Scheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **65**, 810-815 (1970).
27. Lim, V. I. *J. molec. Biol.* **88**, 857-872, 873-894 (1974).
28. Schulz, G. E. *et al. Nature* **250**, 140-142 (1974).
29. Matthews, B. W. *Biochim. biophys. Acta* **405**, 442-451 (1975).
30. Argos, P., Schwarz, J. & Schwarz, J. *Biochim. biophys. Acta* **439**, 261-273 (1976).
31. Richards, F. M. *J. molec. Biol.* **82**, 1-14 (1974).
32. Chothia, C., Levitt, M. & Richardson, D. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4130-4134 (1977).
33. Lee, B. & Richards, F. M. *J. molec. Biol.* **55**, 379-400 (1971).
34. Chothia, C. *Nature* **248**, 338-339 (1974).
35. Kuntz, I. D. *J. Am. chem. Soc.* **94**, 4009-4012 (1972).
36. Chothia, C. *J. molec. Biol.* **75**, 295-302 (1973).
37. Sternberg, M. J. E. & Thornton, J. M. *J. molec. Biol.* **105**, 367-382 (1976); **110**, 269-283 (1977).
38. Richardson, J. S. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2619-2623 (1976).
39. Nagano, K. *J. molec. Biol.* **109**, 235-250 (1977).
40. Phillips, D. C., Sternberg, M. J. E., Thornton, J. M. & Wilson, I. A. *J. molec. Biol.* (in the press).
41. Levitt, M. & Chothia, C. *Nature* **261**, 552-558 (1976).
42. Sternberg, M. J. E. & Thornton, J. M. *J. molec. Biol.* **110**, 285-296 (1977).
43. Richardson, J. S. *Nature* **265**, 495-500 (1977).

44. Sternberg, M. J. E. & Thornton, J. M. *J. molec. Biol.* **115**, 1–17 (1977).
45. Rao, S. T. & Rossmann, M. G. *J. molec. Biol.* **76**, 241–256 (1973).
46. Dickerson, R. E. & Timkovitch, R. *The Enzymes* (ed. Boyer, P. D.) **11**, 397–593 (Academic Press, New York, 1975).
47. Browne, W. J., North, A. C. T., Phillips, D. C., Brew, K., Vanaman, T. C. & Hill, R. C. *J. molec. Biol.* **42**, 65–86 (1969).
48. Kreisinger, R. H. & Barry, C. D. *Biochim. biophys. Acta* **405**, 40–52 (1975).
49. Rossmann, M. G., Liljas, A., Branden, C.-I. & Banaszak, L. J. *The Enzymes* (ed. Boyer, P. D.) **11**, 61–102 (Academic Press, New York, 1975).
50. Stellwagen, E. *J. molec. Biol.* **106**, 903–911 (1976).
51. Wootton, J. C. *Nature* **252**, 542–546 (1974).
52. Levitt, M. & Warshel, A. *Nature* **253**, 694–698 (1975).
53. Kuntz, I. D., Crippen, G. M., Kollman, P. A. & Kimelman, D. *J. molec. Biol.* **106**, 983–994 (1976).
54. Tanaka, S. & Scheraga, H. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1320–1323 (1977).
55. Levitt, M. *J. molec. Biol.* **104**, 59–107 (1976).
56. Pitsyn, O. B. & Rashin, A. A. *Biophys. Chem.* **3**, 1–20 (1975).
57. Lim, V. I. & Efimov, A. V. *FEBS Lett.* **69**, 41–44 (1976).
58. Zav'yalov, V. P. *Biochim. biophys. Acta* **490**, 506–514 (1977).
59. Sternberg, M. J. E. & Thornton, J. M. *J. molec. Biol.* **113**, 401–418 (1977).

articles

SU(2) × U(1): A gauge theory of weak interactions?

David Bailin & Norman Dombey

School of Mathematical and Physical Sciences, University of Sussex, Falmer, Brighton, Sussex, UK

We construct a gauge theory of weak interactions which fits all neutral current data, including the atomic physics experiments. It is shown that it is possible to include this theory in a unified theory of weak and electromagnetic interactions; the desirability of doing so is discussed.

In 1967 Weinberg¹ wrote down a renormalisable unified gauge theory of weak and electromagnetic interactions based on the spontaneously broken gauge group SU(2) × U(1). This model involves the existence of a weak neutral current coupled to neutrinos, and weak neutral currents were discovered experimentally² in neutrino scattering in 1973. So it has been a great temptation for physicists, especially theorists, to claim that the discovery of neutral currents is a confirmation of the Weinberg model and in particular shows the unification of weak and electromagnetic interactions.

This view although widespread has been forcefully criticised, especially by Sakurai³. He argues that neutral current phenomena are interesting in their own right, and that it is the job of physicists to discover through experiment the relevant couplings not to assume the results in advance. Nevertheless the Weinberg model has been very successful and all neutrino data up to now are consistent with a single value of $\sin^2\theta_w$. The model is obviously useful as it not only provides a means of analysing experiments but it also suggests new experiments. In Bjorken's words⁴ it is a "working hypothesis". In particular in its simplest form it gives definite predictions of parity violation in heavy atoms which were not observed in recent experiments at Oxford and Washington⁵.

The purpose of this article is to point out another possibility. The starting point is a remark made some years ago⁶: the requirement that the theory of weak interactions is renormalisable has nothing *a priori* to do with unification of weak and electromagnetic interactions. After all it is (or at least was) customary to treat strong, electromagnetic and weak interactions independently of each other.

We shall treat this remark perhaps more seriously than it deserves in order to construct a new working hypothesis. We shall write down a SU(2) × U(1) gauge theory of weak interactions which as far as neutrino interactions are concerned allows an almost identical analysis of the present data. This theory gives parity violation in heavy atoms but the magnitude of the effect predicted is now in agreement with experiment. The theory is renormalisable in the absence of electromagnetic and strong interactions.

We shall construct our weak SU(2) × U(1) model first and discuss its difficulties afterwards. We start with the leptonic

sector. Following Weinberg we take a doublet $L = (\nu, e)_L$ of left-handed leptons and a right-handed lepton singlet $R = e_R$ with a Lagrangian

$$\mathcal{L} = i\bar{L}\gamma^\mu(\partial_\mu + \frac{1}{2}ig\tau\cdot\mathbf{W}_\mu - \frac{1}{2}ig'B_\mu)L + i\bar{R}\gamma^\mu(\partial_\mu - irB_\mu)R + \mathcal{L} \quad (1)$$

where $\mathcal{L}$ involves the meson-meson couplings required by gauge invariance.

Take

$$\begin{aligned} Z_\mu &= W_\mu^3 \cos\theta_w - B_\mu \sin\theta_w \\ X_\mu &= W_\mu^3 \sin\theta_w + B_\mu \cos\theta_w. \end{aligned} \quad (2)$$

In the unified theory X^0 is the photon and the constraints are (1) no $\bar{\nu}\nu X^0$ coupling and (2) the $\bar{e}eX^0$ coupling is pure vector. These conditions give $g' = g \tan\theta_w$ and $r = g'$. This leads to a Lagrangian

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_0 - \frac{g}{2\sqrt{2}}[\bar{\nu}\gamma^\alpha(1-\gamma_5)eW_\alpha^+ + h.c.] \\ &- \frac{g'}{4\sin\theta_w}Z^\alpha[\bar{\nu}\gamma_\alpha(1-\gamma_5)\nu - (1-4\sin^2\theta_w)\bar{e}\gamma_\alpha e + \bar{e}\gamma_\alpha\gamma_5 e] \\ &+ g\sin\theta_w X^\alpha \bar{e}\gamma_\alpha e + \mathcal{L}. \end{aligned} \quad (3)$$

In weak SU(2) × U(1) X^0 is not the photon but a second neutral vector meson. The requirement that there should be the same relation between neutrino scattering by charged currents and neutral currents in this model as in the unified theory implies again that there is no $\bar{\nu}\nu X^0$ coupling. So again $g' = g \tan\theta_w$. In order to obtain a distinct alternative theory we require now that the $\bar{e}eX^0$ coupling is pure axial. This gives $r = -g'$. With these choices the Lagrangian becomes

$$\begin{aligned} \mathcal{L} &= \mathcal{L}_0 - \frac{g}{2\sqrt{2}}[\bar{\nu}\gamma^\alpha(1-\gamma_5)eW_\alpha^+ + h.c.] \\ &- \frac{g'}{4\sin\theta_w}Z^\alpha[\bar{\nu}\gamma_\alpha(1-\gamma_5)\nu + (1-4\sin^2\theta_w)\bar{e}\gamma_\alpha\gamma_5 e - \bar{e}\gamma_\alpha e] \\ &- g\sin\theta_w X^\alpha \bar{e}\gamma_\alpha\gamma_5 e + \mathcal{L}. \end{aligned} \quad (4)$$

Comparing the Lagrangians (3) and (4) we see that there is no difference in neutrino couplings but that in the $\bar{e}eZ^0$ couplings the vector and axial currents are interchanged.

The GIM⁷ mechanism can now be used as in the unified theory to construct the weak hadronic interaction. We will not display

the result; all we need is that whereas in unified $SU(2) \times U(1)$ the $\Delta S = \Delta C = 0$ neutral weak hadronic current is

$$J_\mu = A_\mu^3 + (1 - 2\sin^2\theta_w)V_\mu^3 - 2\sin^2\theta_w V_\mu^8, \quad (5)$$

in weak $SU(2) \times U(1)$ the current coupling to Z^0 is

$$J_\mu = V_\mu^3 + (1 - 2\sin^2\theta_w)A_\mu^3 - 2\sin^2\theta_w A_\mu^8, \quad (6)$$

where V_μ and A_μ are the vector and axial vector currents, and 3 and 8 refer to the isovector and isoscalar components. There is also a purely axial current which couples to X^0 which will not contribute to neutrino scattering.

Before turning to the neutral current data, we must define the strength of neutral current neutrino processes compared with charged current processes in the effective current theory. The Fermi constant G for charged currents is given in terms of the $W^\pm$ mass m_w by

$$\frac{G}{\sqrt{2}} = \frac{g^2}{8m_w^2}. \quad (7)$$

Similarly, if the Fermi constant for neutral currents is G' , then from equation (4)

$$\frac{G'}{\sqrt{2}} = \frac{g'^2}{16m_z^2 \sin^2\theta_w} = \frac{g^2}{16m_w^2 \kappa^2} \quad (8)$$

where $\kappa = m_z \cos\theta_w / m_w$. In the Weinberg model $\kappa = 1$. (Note that Salam and Ward⁸, following Glashow⁹, wrote down an $SU(2) \times U(1)$ unified gauge theory in 1964. As it was not a spontaneously broken gauge theory, however, it was not renormalisable and there were no predictions of the vector mesons' masses. Thus the structure of their weak neutral current is that given by Weinberg, but the relative strength of the neutral current compared with the charged current in neutrino interactions is not specified. Only with a detailed symmetry breaking mechanism can κ be predicted and in our opinion this is the distinguishing feature of Weinberg's model.)

Experimentally $\kappa^2 = 0.97 \pm 0.13$ (ref. 10). In weak $SU(2) \times U(1)$ we will show shortly that $\kappa < 1$, but for the moment we assume that the value predicted is in agreement with experiment. We can now analyse the neutral current data.

Inclusive neutrino reactions

The well-known vector-axial ambiguity¹¹ implies that the results in our model are identical with those of the unified $SU(2) \times U(1)$ model. Thus we can use the Weinberg angle measured in these experiments. The most recent values¹² are $\sin^2\theta_w = 0.31 \pm 0.03$ or $\sin^2\theta_w = 0.26 \pm 0.02$, depending on whether the CDHS (CERN-Dortmund-Heidelberg-Saclay) data is included. We shall be conservative and take $0.24 \leq \sin^2\theta_w \leq 0.34$.

Neutrino-lepton scattering

Again, there is no effect when vector and axial currents are interchanged. So the results of unified $SU(2) \times U(1)$ model are true in weak $SU(2) \times U(1)$.

Elastic νp scattering

The cross section in exclusive channels is not invariant under the interchange of vector and axial vector currents. For example, in the forward direction we have in the Weinberg model from equation (5)

$$R = \lim_{q^2 \rightarrow 0} \left[\frac{d\sigma}{dq^2}(\nu p \rightarrow \nu p) / \frac{d\sigma}{dq^2}(\bar{\nu} n \rightarrow \mu^- p) \right] = \frac{1}{4}[(1 - 4\sin^2\theta_w)^2 + (g_A^3)^2]/[1 + (g_A^3)^2], \quad (9)$$

where $g_A^3 = -1.25$ is the isovector axial renormalisation factor. In weak $SU(2) \times U(1)$ we have from equation (6)

$$R = \frac{1}{4\kappa^2} [1 + \{(1 - 2\sin^2\theta_w)g_A^3 - 2\sin^2\theta_w g_A^8\}^2 / [1 + (g_A^3)^2]]. \quad (10)$$

The isoscalar axial renormalisation factor g_A^8 is not experimentally known, but a reasonably reliable quark model estimate¹³ is $g_A^8 = \frac{2}{3}g_A^3$. Putting in the values for the parameters we obtain in the Weinberg model $0.15 < R < 0.17$, whereas in weak $SU(2) \times U(1)$ $R = 0.10/\kappa^2$ and the dependence on g_A^8 is not very sensitive.

The data are not yet good enough to determine more than total cross sections integrated over all q^2 . For this assumptions must be made about axial form factor behaviour which are not well known, and it is safe to say that any conclusion drawn on the basis of the Weinberg model can be matched in weak $SU(2) \times U(1)$.

Single pion production

This process is very difficult to analyse as most of the data come from heavy liquid bubble chambers where nuclear charge exchange effects are important. There is data from Argonne¹⁴ in hydrogen and deuterium but the errors are still very large. So again it is unlikely that any prediction could be made on the basis of unified $SU(2) \times U(1)$ which could not be reproduced in weak $SU(2) \times U(1)$.

Comparing equations (5) and (6) we see that A_μ^3 is suppressed in weak $SU(2) \times U(1)$ compared with in the Weinberg model, whereas V_μ^3 is enhanced. As A_μ^3 gives the dominant transition in the Δ -resonance region, a calculation of the cross section for a given value of $\sin^2\theta_w$ would give a lower value in weak $SU(2) \times U(1)$ than in unified $SU(2) \times U(1)$. The theoretical analysis¹⁵ of weak pion production in unified $SU(2) \times U(1)$ shows that cross sections decrease as $\sin^2\theta_w$ increases. Hence if weak $SU(2) \times U(1)$ is right an analysis of this process in unified $SU(2) \times U(1)$ would give a relatively large value of $\sin^2\theta_w$.

Parity violation in atoms

Z^0 exchange between an atomic electron and the nucleus will not conserve parity. This effect arises both from a vector interaction at the electron together with an axial interaction at the nucleus, and from an axial interaction at the electron together with a vector interaction at the nucleus. Taking the nucleus as static the latter term will dominate in heavy nuclei as it gives a coherent contribution over the whole nucleus. In the Weinberg model the coefficient of the axial coupling is taken as unity from equation (3) whereas the nuclear vector contribution from equation (5) gives an overall effect

$$Q_w = (1 - 4\sin^2\theta_w)Z - N. \quad (11)$$

Numerically in bismuth $123 < -Q_w < 156$. From the experiment⁵, the observed value of Q_w in bismuth is no larger than one-tenth of the maximum value predicted by the Weinberg model together with the standard atomic theory. The standard atomic theory should be good to at least a factor two¹⁶, so it is fair to say that $|Q_w| < 32$ for bismuth.

In weak $SU(2) \times U(1)$ we have a factor $1 - 4\sin^2\theta_w$ for the axial coupling from equation (4), and the nuclear vector coupling is pure isovector from equation (6). The exchange of X^0 conserves parity. So now

$$Q_w = (1 - 4\sin^2\theta_w)(Z - N)/\kappa^2 \quad (12)$$

or in bismuth $-2 < Q_w \kappa^2 < 15$, which is not in disagreement with the experiments.

If $\sin^2\theta_w = \frac{1}{4}$ then $Q_w = 0$, as the $\bar{e}eX^0$ coupling is now pure vector. This would rule out any coherent effect although there would still be parity violation from the axial coupling of the nucleus. This would be seen most clearly in experiments on

hydrogen or deuterium or in the experiment to measure a macroscopic electric dipole moment in superfluid Helium 3 suggested by Leggett¹⁷. It is also worth noting that for $\sin^2\theta_w = \frac{1}{4}$ in weak $SU(2) \times U(1)$ there is no parity violation in purely leptonic neutral current interactions. So there is no forward-backward asymmetry in $e^+e^- \rightarrow \mu^+\mu^-$ and $\sigma(\nu_\mu e^-) = \sigma(\bar{\nu}_\mu e^-)$.

We now turn to the mass relations of the gauge particles in weak $SU(2) \times U(1)$. We require spontaneous breaking with respect to $SU(2)$ through a single isospinor complex scalar field as in the Weinberg theory, but now we also spontaneously break the $U(1)$ through an isoscalar complex scalar field, as the current coupled to the X^0 is no longer conserved. Thus two Higgs scalars result and the particles $W^\pm$, X^0 , Z^0 acquire mass. The mass relations follow from the mixing equations

$$\begin{aligned} m_x^2 &= m_w^2 \sin^2\theta_w + m_B^2 \cos^2\theta_w + 2m_{WB}^2 \sin\theta_w \cos\theta_w \\ m_z^2 &= m_w^2 \cos^2\theta_w + m_B^2 \sin^2\theta_w - 2m_{WB}^2 \sin\theta_w \cos\theta_w \\ 0 &= m_{WB}^2 (\cos^2\theta_w - \sin^2\theta_w) + (m_w^2 - m_B^2) \sin\theta_w \cos\theta_w \end{aligned} \quad (13)$$

where m_{WB}^2 is the off-diagonal element of the mass matrix. So

$$m_w^2 = m_z^2 \cos^2\theta_w + m_x^2 \sin^2\theta_w. \quad (14)$$

For $m_x = 0$ this reduces to $m_w = m_z \cos\theta_w$ and $\kappa = 1$, as in the Weinberg model. Now, however, we are constrained by the experimental value of κ . Since $\kappa \sim 1$, m_x cannot be too large. Nevertheless, since the data allow $\kappa^2 > 0.84$, the previously given bounds on $\sin^2\theta_w$ enable us to deduce from equation (14) that $m_x/m_w < 0.7-0.8$.

Note that there is no difficulty in giving mass to X^0 , so it cannot be claimed that a photon is a necessary concomitant of the theory. If the gauge group were simple, as in the Georgi-Glashow model¹⁸, and if the symmetry breaking were through the regular representation¹⁹, then it would be natural to have a zero mass vector meson in the theory: theories of this type are discussed elsewhere²⁰. In weak $SU(2) \times U(1)$ the value of m_w is not constrained. In unified $SU(2) \times U(1)$ $g \sin\theta_w = e$ and using equation (7) this fixes the mass scale of the theory so that $m_w > 37$ GeV. In weak $SU(2) \times U(1)$ g is just the weak coupling constant and is arbitrary. Indeed this model is flexible: if no vector mesons are found then the limit $g \rightarrow \infty$, $m_w \rightarrow \infty$ can be taken and the old four-fermion theory reappears.

It is now time to mention the difficulty of a gauge theory of weak interactions which leaves out electromagnetism. The process $e^+e^- \rightarrow W^+W^-$ through one photon exchange displays all the old problems associated with charged spin on particles. In general the electromagnetic corrections to the theory are not renormalisable.

Of course the obvious way out of this difficulty is to reunify the theory after all by including the photon as a gauge particle and enlarging the group. The minimal extension of weak $SU(2) \times U(1)$ is $SU(2) \times U(1) \times U(1)$, where the new $U(1)$ is characterised by another isoscalar vector meson C_μ and the photon field

$$a_\mu = W_\mu^3 \sin\psi + C_\mu \cos\psi. \quad (15)$$

Then, as the resultant theory is unified, $e = g \sin\psi$, and immediately we obtain the unified mass scale $m_w = 37/|\sin\psi|$ GeV. Provided $\sin\psi$ is reasonably small all the results of weak $SU(2) \times U(1)$ will still hold, up to an error of order $|\sin\psi|$. There is now no connection between $\sin\theta_w$ determined from neutrino scattering and m_w . Equation (14) still holds approximately and so although $W^\pm$ must now have a large mass X^0 can still be light²¹.

The historic reasons^{8,9} for a non-Abelian gauge theory of weak interactions are that these involve charged currents which are conserved—the natural way of ensuring this is to construct a gauge theory. Alternative possibilities to a unified theory of weak and electromagnetic interactions are suggested by analogy with strong interaction theory where there were also good reasons for writing a non-Abelian gauge theory in order to explain isotopic spin conservation. Sakurai²² introduced the ρ meson as a gauge particle for this purpose. Gell-Mann²³ constructed the $SU(3)$

symmetry of strong interactions by generalising these ideas. Yet these theories involved charged vector mesons $\rho^\pm$ and therefore were open to the same criticism as here: these theories are non-renormalisable in the presence of electromagnetism.

These criticisms are no longer relevant. The modern picture of $\rho^\pm$ is a composite of two spin- $\frac{1}{2}$ particles, a quark-antiquark pair. The ultra-high energy behaviour of $e^+e^- \rightarrow W^+W^-$ similarly depends on the internal electromagnetic structure of $W^\pm$. Form factors arising from a composite nature (a pair of quarks, heavy leptons, leptoquarks?) or a hidden strong interaction of $W^\pm$ could save the day. So could a reinterpretation of quantum electrodynamics for high spin particles along the lines suggested by Lee and Yang²⁴.

All gauge theories require the existence of intermediate vector mesons. In a unified gauge theory the mass scale is fixed: $m_w \sim e/G^{1/2} = 88$ GeV. In a weak gauge theory this is not necessarily so, and, if weak interactions were really weak at all energies, we would expect relatively light $W^\pm$, for example if $g^2/4\pi = 10^{-3}$, then $m_w = 14$ GeV. Experimentally the lower limit on m_w is 15 GeV (D. H. Perkins, personal communication).

In weak $SU(2) \times U(1)$ a neutral vector meson X^0 exists which must be lighter than $W^\pm$. It has not yet been observed in e^+e^- colliding beam experiments at SPEAR, so $m_x > 7.6$ GeV (ref. 25). It could be identified by its (relatively) narrow width

$$\Gamma = (2+R) \frac{G\sqrt{2}}{3\pi} m_w^2 m_x \sin^2\theta_w \quad (16)$$

where in this model R is the cross section ratio $\sigma(e^+e^- \rightarrow \text{hadrons})/\sigma(e^+e^- \rightarrow \mu^+\mu^-)$ at centre-of mass energy m_x . So if $m_x = 10$ GeV and $m_w = 90$ GeV, then $\Gamma \sim 250$ MeV.

We should emphasise that even if the theory is unified, so that $SU(2) \times U(1) \times U(1)$ or a larger group such as $SU(2) \times SU(2) \times U(1)$ (refs 26, 27) or $SU(3) \times U(1)$ (ref. 28) is applicable, it may still be useful to consider weak $SU(2) \times U(1)$ as an approximate symmetry for weak interactions. Indeed the Weinberg model is perhaps too constrained a theory. In weak $SU(2) \times U(1)$ $\sin\theta_w$ fixes the form of the neutral currents but it does not fix the strength relative to the charged current; for that κ is needed, where $\kappa \neq 1$. Again, in the Weinberg model the neutral vector meson must be heavier than $W^\pm$, while in weak $SU(2) \times U(1)$ the X^0 can be arbitrarily light. Only if and when the vector mesons are found can definite statements be made about the underlying theory. It follows that experimentalists should continue to look for $W^\pm$ and neutral weak vector mesons, and not assume that the success of the Weinberg model implies that they can only be found with a new generation of accelerators.

Finally we point out that there is one good (if not compelling) reason for the unification of weak and electromagnetic interactions which results from experiment. This comes about by inverting the argument of Wilkinson²⁹ on the radiative corrections to vector current universality. If the observed muon decay rate is compared with nuclear ft-values, it is found that

$$\Delta \equiv 2\{g_{\beta\nu}^R/g_\mu^R \cos\theta_c - 1\} = 2.19 \pm 0.27\% \quad (17)$$

where $g_{\beta\nu}^R$ and g_μ^R are the effective β decay and muon decay coupling constants (after removing the well-understood 'outer' radiative corrections) and θ_c is the Cabibbo angle. Theoretically the purely electromagnetic contribution to Δ , arising from the 'inner' (short distance) radiative corrections gives a divergent contribution³⁰

$$\Delta = \frac{3\alpha}{2\pi} (1 + 2\bar{Q})(\ln(\Lambda/m_N)) \quad (18)$$

where Λ is a cut-off and $\bar{Q}$ is the average charge of the quark constituents of the nucleon. Assuming $\bar{Q} = \frac{1}{6}$, as in appropriate for both the conventional Gell-Mann Zweig quarks used throughout this analysis and integrally charged Han Nambu quarks³¹, we obtain from equations (16) and (17) that $\Lambda \approx 90$ GeV.

In a unified theory of weak and electromagnetic interactions the average mass of the gauge vector mesons would provide an

effective cut-off³²⁻³⁴. So we conclude that this result implies that the value of m_w is about 90 GeV (ref. 35), which is consistent with the mass scale expected in a unified theory.

We thank Roger Blin-Stoyle, Alex Love, Don Perkins and Patrick Sandars for discussions.

Received 9 October; accepted 30 October.

- Weinberg, S. *Phys. Rev. Lett.* **19**, 1264 (1967).
- Hasert, F. J. *et al. Phys. Lett.* **46B**, 138 (1973); *Nucl. Phys.* **B73**, 1 (1974).
- Sakurai, J. J. *Lecture notes in Physics* **56**, 258 (Springer, Berlin 1976).
- Bjorken, J. D. in *Weak Interactions at High Energy and the Production of New Particles* (ed. Zif, M. C.) I SLAC-198 Stanford 1976.
- Baird, P. E. G. *et al. Nature* **264**, 528 (1976); *Phys. Rev. Lett.* **39**, 795, 798 (1977).
- Dombey, N. *Lectures on Neutrino scattering and gauge models* University College London Internal Report (1974).
- Glashow, S., Iliopoulos, J. & Maiani, L. *Phys. Rev. D* **2**, 1285 (1970).
- Salam, A. & Ward, J. C. *Phys. Lett.* **13**, 168 (1964).
- Glashow, S. *Nucl. Phys.* **22**, 579 (1961).
- Barger, V. & Nanopoulos, D. V. *Nucl. Phys.* **B24**, 426 (1977).
- Hung, P. Q. & Sakurai, J. J. *Phys. Lett.* **63B**, 295 (1976).
- Musset, P. presented at *Int. Symp. Lepton and Photon Interactions at High Energies Hamburg* (August 1977).

- Adler, S. L. *et al. Phys. Rev. D* **11**, 3309 (1975).
- Hyman, L. G. *Argonne Rep.* ANL-HEP-PR-75-08 (1975).
- Alevisos, T. & Dombey, N. J. *Phys. G* **3**, L43 (1977).
- Sandars, P. G. H. presented at *Int. Symp. Lepton Photon Interactions at High Energies Hamburg* (August 1977).
- Leggett, A. J. *Phys. Rev. Lett.* **39**, 587 (1977).
- Georgi, H. & Glashow, S. L. *Phys. Rev. Lett.* **28**, 1494 (1972).
- Higgs, P. W. *Phys. Rev. Lett.* **13**, 508 (1964).
- Dombey, N. & Vayonakis, C. E. J. *Phys. A* **9**, 1381 (1976).
- Elias, V. & Swift, A. R. *Phys. Rev. D* **15**, 1937 (1977).
- Sakurai, J. J. *Annals Phys. N.Y.* **11**, 1 (1960).
- Gell-Mann, M. *Cal. Inst. Technol. Rep. CTS-20* (1961).
- Lee, T. D. & Yang, C. N. *Phys. Rev.* **128**, 885 (1962).
- Schwitters, R. F. *Proc. 1975 Int. Symp. Lepton Photon Interactions at High Energies* (ed. Kirk, W. T.), SLAC Stanford 1976, p. 5.
- Fritsch, H. & Minkowski, P. *Nucl. Phys.* **B103**, 61 (1976).
- Mohapatra, R. & Sidhu, D. P. *Phys. Rev. Lett.* **38**, 667 (1977).
- Lee, B. W. & Weinberg, S. *Fermilab Publ.* 77/31-THY (1977).
- Wilkinson, D. H. *Nature* **257**, 189 (1975); *Phys. Lett.* **67B**, 13 (1977).
- Bjorken, J. D. *Phys. Rev.* **148**, 1467 (1966); *E160*, 1582 (1967).
- Sirlin, A. *Nucl. Phys.* **B100**, 291 (1975).
- Lee, T. D. *Phys. Rev.* **128**, 899 (1962).
- Bailin, D. *Phys. Rev.* **135B**, 166 (1964).
- Mohapatra, R. N. & Vinciguelli, P. *Phys. Rev.* **D8**, 481 (1973).
- Blin-Stoyle, R. J. *Fundamental Interactions and the Nucleus*, 51 (North Holland, Amsterdam 1972).

Granite-greenstone terrains of the Rhodesian Archaean craton

J. F. Wilson

Department of Geology, University of Rhodesia, Salisbury, Rhodesia

M. J. Bickle, C. J. Hawkesworth

Department of Earth Sciences, University of Leeds, Leeds, UK

A. Martin

Department of Geology and Mineralogy, Parks Rd, Oxford, UK

E. G. Nisbet

Institut für Kristallographie und Petrographie, ETH-Zentrum, CH-8092, Zürich, Switzerland

J. L. Orpen

Department of Geology, University of Rhodesia, Salisbury, Rhodesia

Granite-greenstone terrains of possibly three different ages exist in the Rhodesian Archaean craton. Their respective greenstone belts can be tentatively delineated across the central cratonic area. The main stratigraphic units of the youngest (~ 2,700 Myr) and most widespread development can be correlated across greenstone belts in this area. Major events in the craton are summarised and briefly discussed in the light of the new correlations and recent age determinations.

A REVIEW¹ of available field and geochronological data allows the recognition and partial delineation of granite-greenstone terrains of different ages within the Basement Complex of the Rhodesian Archaean craton. The oldest terrain is ~ 3,500 Myr old and occurs in the southern and central parts (Fig. 1 and Table 1). Its greenstone belt remnants equate largely with the Sebakwian Group of the Rhodesia Geological Survey as delineated by Wiles². The main greenstone belts are younger and can be divided into the widespread Upper Greenstones and the more problematical Lower Greenstones which together correspond largely to the Bulawayan Group of Wiles. The Sebakwian and Bulawayan Groups as delineated by Anhaeusser³, on the basis of the Barberton model, do not correspond to those of Wiles, or to the subdivisions recognised in this paper.

The age of the Lower Greenstones is obscure. They were deformed and eroded before the deposition of the Upper Greenstones. The results obtained so far on Lower Greenstones material yield ages in the ~ 2,700 Myr bracket, but initial ⁸⁷Sr/⁸⁶Sr ratios are higher than the near 0.701 values typical of the Upper Greenstones. Whether such high initial ratios (for example 0.7056 ± 0.002, Hokonui tuffs⁴) merely reflect the primary composition of these rocks, or can be taken as evidence for metamorphic rotation of the isochron to an age significantly younger than their age of formation, is still debatable. On

regional grounds¹ it is possible that the Lower Greenstones are part of a granite-greenstone terrain incorporating tonalites and gneisses (for example the Mashaba tonalite and Chingezi gneisses south-west of Belingwe) dated at ~ 2,900 Myr (Table 1; refs 4,5). The Upper Greenstones are ~ 2,700-Myr-old and the low initial ⁸⁷Sr/⁸⁶Sr ratios are consistent with this being an age of eruption⁶. In the south-east and the west the Upper Greenstones are overlain by the sedimentary Shamvaian Group.

Jahn and Condie⁷ have recently offered an age of 3,080 ± 60 Myr (2σ errors) for the Bulawayo-Que Que greenstone belt, but it is difficult to place any geological significance on this result. Geographically their data points involve a spread of 200 km and stratigraphically they range from Lower Greenstones to the upper part of the Upper Greenstones. Moreover their selection of 12 data points from the 50 considered, seems somewhat arbitrary^{1,4}.

The ~ 3,500 Myr terrain

The largest development of ~ 3,500-Myr-old rocks occurs within a roughly triangular crustal segment with Selukwe, Fort Victoria and Shabani near the approximate corners⁸. Much of this segment consists of gneisses such as those around Shabani, and the Tokwe gneisses west of Mashaba (Table 1). Most are tonalitic and highly deformed and, together with the infolded greenstone belt remnants, look like the remains of an ancient, northerly trending, mobile belt. Between Mashaba and Fort Victoria the gneisses are cut by the ~ 3,500-Myr-old Mushandike granite^{9,10}, and elsewhere by younger granitic rocks¹. The greenstone belt remnants within the segment, and the presumed ~ 3,500-Myr-old remnants elsewhere in the central cratonic area (Fig. 1), consist of various sedimentary and igneous metasedimentary rocks and other mafic and ultra-mafic types. Felsic metavolcanic rocks are apparently absent, although they would be difficult to recognise within the gneisses. Outside the triangular segment it is not clear how much of the associated gneissic terrain is ~ 3,500 Myr old.

The best documented of the larger areas of ~ 3,500-Myr-old

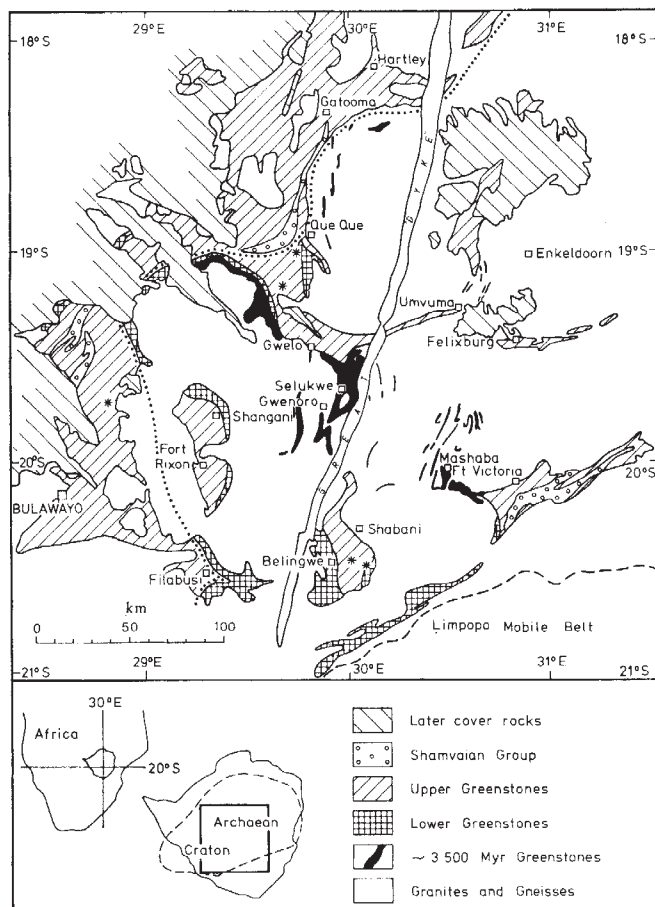


Fig. 1 Subdivisions of greenstone belts in central area of Rhodesian Archaean craton. *, Stromatolites. Dotted line represents approximate division between western and eastern successions of Upper Greenstones.

greenstone belts is the Selukwe area. Here the succession is inverted and forms the lower limb of a large recumbent fold structure, called the Selukwe nappe¹¹⁻¹⁵. This structure is cut by granite of the Mont d'Or formation dated at ~3,400-Myr-old¹⁶. The rocks of the Selukwe greenstone belt are largely metavolcanic, ranging from basaltic to peridotitic, but they also include a chromitite-bearing ultramafic intrusion. About half-way up the succession is the sedimentary Wanderer formation, the basal conglomerate of which rests with a marked unconformity on the deformed and eroded ultramafic intrusion. Boulders within this conglomerate are particularly significant since they include gneissic tonalite and granodiorite, massive adamellite and pegmatite¹⁵, indicating the presence of complex sialic crust predating at least part of the ~3,500-Myr-old greenstone belts.

The ~2,900 Myr terrain

A major ~2,900-Myr-old granitic event, which was considered to represent a minimum age for the main greenstone belts, was postulated by Wilson¹⁷ on the basis of conventional K/Ar results¹⁸ from a number of granites and gneisses in the central cratonic area. Subsequent Rb/Sr work disproved this and showed these K/Ar data to be quite unreliable as indicators of age of intrusion.

Two more recent Rb/Sr results^{4,5}, however, confirm the presence of ~2,900-Myr-old granitic rocks within the craton. These rocks, which form part of the pre-Upper Greenstones basement, are the homogeneous, slightly foliated Mashaba tonalite, which effectively separates the ~3,500-Myr-old rocks between Mashaba and Selukwe, and the highly flattened tonalitic Chingezi gneisses¹⁹ on the west side of the Belingwe greenstone belt (Table 1). The Chingezi tonalite, lithologically similar to the Mashaba tonalite, intrudes the Chingezi gneisses and the Lower

Greenstones (Hokonui formation) west of Belingwe. It is so far undated, but it may also be part of the ~2,900-Myr-old terrain.

The main greenstone belts

J.F.W.¹ has correlated the main stratigraphic units of the Belingwe greenstone belt²⁰⁻²³ across the central cratonic area. The Belingwe belt consists of two sequences, the Lower and Upper Greenstones, separated by an unconformity and by at least one period of deformation. The Lower Greenstones flank the Upper Greenstones on the western and southeastern sides of the belt (Fig. 1).

The topmost (Bend) formation, in the western development of Belingwe Lower Greenstones, consists of alternating pillowed mafic lavas, spinifex-textured peridotites and intercalated banded iron-formation, capped by a thick conglomerate member which contains an interbedded unit of felsic agglomerate. Boulders in the conglomerate include granitic types. Underlying the Bend formation is the Hokonui formation (Table 1) which consists predominantly of felsic flows and pyroclasts. Below the Hokonui formation is a poorly exposed unit of amphibolites, infolded with the ~2,900-Myr-old Chingezi gneisses. The status of this unit is obscure, but we have included it with the Lower Greenstones in Fig. 1.

The Lower Greenstones on the south-east side of the Belingwe belt (the Brooklands formation) pass upwards from quartzites, chloritic grits and conglomerates into high-magnesium lavas which are locally pillowed and alternate with phyllites and banded iron-formation. The conglomerates contain tonalite clasts. The basal contact of the Brooklands formation is not exposed, but its relatively undeformed greenschist facies rocks abut on to complexly deformed ~3,500-Myr-old gneisses⁵.

The thin Manjeri formation forms the base of the unconformably overlying Belingwe Upper Greenstones. It consists of local developments of clastic shallow water sediments and limestones, capped by a persistent horizon of predominantly sulphide facies banded iron-formation. On the west side of the belt the Manjeri formation lies on Lower Greenstones; on the east side, it transgresses the Brooklands formation to lie with well-exposed unconformity on the ~3,500-Myr-old Shabani gneisses^{20,21}. Succeeding the Manjeri formation is a thick tholeiitic volcanic pile of massive and pillowed basalt and metabasalt flows with

Table 1 Rb/Sr whole rock isochron ages and initial ⁸⁷Sr/⁸⁶Sr ratios from the Rhodesian Archaean craton

Rock unit	Age (Myr)*	⁸⁷ Sr/ ⁸⁶ Sr initial
Great Dyke ^{4,7}	2,514 ± 16	0.7026 ± 0.0004
Late Granites-Chilimanzi Suite		
Chilimanzi ^{10,36}	2,620 ± 80	0.7035 ± 0.0080
Zimbabwe ^{10,36}	2,610 ± 60	0.7044 ± 0.0018
Victoria Porphyritic ^{10,36}	2,660 ± 70	0.7025 ± 0.0030
All three on one isochron ³⁶	2,625 ± 25	0.7040 ± 0.0010
Late Granites-Sesombi Suite		
Sesombi ⁶	2,690 ± 140	0.7008 ± 0.0008
Somabula ⁸	2,650 ± 80	0.7012 ± 0.0004
Upper Greenstones		
Maliyami Formation, Que Que area ⁶	2,720 ± 140	0.7010 ± 0.0004
Bulawayo greenstone belt ⁶	2,540 ± 180	0.7015 ± 0.0004
Salisbury greenstone belt ⁷	2,730 ± 60	0.7012 ± 0.0007
Mixed Upper and Lower Greenstones		
Belingwe greenstone belt ⁷	2,760 ± 70	0.7029 ± 0.0002
'Lower Bulawayan', Que Que area ⁶	2,530 ± 280	0.7034 ± 0.0012
Lower Greenstones		
Hokonui Formation, Belingwe ^{4,5}	2,620 ± 120	0.7056 ± 0.0005
Gwenoro gneisses, south-west of Selukwe ⁶	2,780 ± 60	0.7011 ± 0.0002
Gneisses, Umniati River, Rhodesdale batholith, north-east of Que Que ⁸	2,760 ± 80	0.7015 ± 0.0004
Mashaba tonalite ^{4,5}	2,970 ± 160	0.7013 ± 0.0008
Chingezi Gneisses, Belingwe ⁵	2,884 ± 92	0.7017 ± 0.0005
~3,500 Myr terrain		
Mont d'Or formation granite, Selukwe ¹⁶	3,420 ± 120	0.711 ± 0.002
Mushandike granite ^{9,36}	3,520 ± 260	0.7017 ± 0.0030
Gneisses, Shabani area ⁸	3,570 ± 120	0.7000 ± 0.0010
Tokwe Gneisses, Mashaba area ⁶	3,580 ± 400	0.701 ± 0.002

*⁸⁷Rb decay constant $1.39 \times 10^{-11} \text{ yr}^{-1}$. All errors at 2σ level.

Table 2 Simplified western and eastern successions of the Upper Greenstones¹

Western succession	Eastern succession
Calc-alkaline suite of basalt, andesite, dacite flows and pyroclasts; some mafic sills. (? ± 4 km thick)	Not developed
Bimodal volcanic suite: tholeiitic and magnesium-rich basalt and metabasalt pillowed and massive flows, some peridotites; alternating with dacitic flows, tuffs and agglomerates. Grits and conglomerates locally derived from pyroclasts. Sericitic schists. Mafic sills; banded iron-formation phyllites, local limestones in places stromatolitic. (? ± 4 km thick)	Tholeiitic and some magnesium-rich basalt and metabasalt pillowed and massive flows. (? 1–2 km thick)
As for eastern succession	Phyllites, banded iron-formation, local conglomerates and grits; some limestones in places stromatolitic. (Up to 2 km thick)
As for eastern succession	Tholeiitic basalt and metabasalt pillowed and massive flows. Basal development of peridotitic and magnesium-rich basalt flows. Subsidiary sills. Minor pyroclasts. (Up to 7 km thick)
As for eastern succession	Banded iron-formation, local conglomerates and grits, some limestones in places stromatolitic. (Up to 0.25 km thick)

some high-magnesium basalt and subsidiary sills, the whole largely devoid of sediments. The lowermost part of this volcanic sequence is dominated by high-magnesium rocks including pillowed and spinifex-textured peridotite flows and is known as the Reliance formation; the thicker, major part is the Zeederbergs formation^{20,24}. Capping the Zeederbergs formation are the dominantly pelitic shallow-water sedimentary rocks of the Cheshire formation, with locally developed banded iron-formation and limestone.

Figure 1 shows the distribution of the Lower and Upper Greenstones in the central cratonic area¹. Detailed comparisons within the Lower Greenstones are difficult but elements of the Belingwe succession can be recognised. Prominent in all occurrences west of the Great 'Dyke' is a thick unit of Hokonui-type felsic volcanic rocks and derived schists, whereas the linear Mweza greenstone belt²⁵, south-east of Belingwe, contains a sequence possibly transitional between the Bend and Brooklands formations. North-west of Gwelo the Lower Greenstones lie with marked unconformity (refs 1, 26 and D. Edwards personal communication) on presumed ~ 3,500-Myr-old greenstones; elsewhere their lowest preserved sequences are intruded by granitic rocks.

Correlations¹ for the Upper Greenstones in the central cratonic area are, however, clearer. In much of the southwestern part of the area, covered by Fig. 1, the base of the Upper Greenstones is indicated by a thin Manjeri-type marker unit which can be traced for many kilometres and from greenstone belt to greenstone belt. This is succeeded by a thick Zeederbergs-type volcanic pile which is conspicuous in all the major greenstone belts and equates with the basaltic greenstones of Macgregor's Bulawayan System^{27,28}. In several areas the basal part of this succession is a high-magnesium Reliance-type volcanic assemblage. Thereafter, however, above this total volcanic pile the succession in the eastern central cratonic area differs from that in the west¹.

In the east the volcanic pile is capped by Cheshire-type sediments, which are overlain by, and in part interbedded with, a further development of tholeiitic lavas not seen at Belingwe. In the west the thick basaltic pile passes up into repeated bimodal mafic-felsic volcanism on different scales in which tholeiitic and in places high-magnesium spinifex-textured (basaltic and peridotitic) flows alternate with felsic flows and pyroclasts. Sediments derived from the felsic rocks, as well as banded iron-formation and a few limestones, are also a feature of this bimodal succession. Where metamorphosed and deformed, the felsic rocks are sericitic schists. Wilson¹ has shown that the western bimodal succession is the time equivalent of, and can be traced laterally into, the eastern pelite-dominant sedimentary unit and its overlying basalts. Suc-

ceeding the bimodal pile is an andesite-dominant, calc-alkaline volcanic assemblage, ranging from basalt to dacite, which was not developed in the east. These two upper western units are confined to a broad zone trending approximately north-north-east across the central cratonic area. In the Que Que and Gatooma areas the bimodal succession includes most of the loosely defined Felsic formation of Harrison^{29,30} and Bliss³¹, whereas the calc-alkaline assemblage embraces the Maliyami formation (Table 1) and the Umniati group respectively of these two authors.

J.F.W.'s¹ correlations across the greenstone belts also allow the rationalisation of the known occurrences of Rhodesian Archean stromatolites which now number five (Fig. 1). All of these occur in the Upper Greenstones. Two occur in limestones of the basal formation below the thick tholeiitic pile; one at Belingwe²⁰, and one at Redcliff south of Que Que (C. A. Castelin, personal communication). Three occur in limestones above the thick tholeiitic pile; one at Belingwe²⁰, one south-east of Redcliff (C. A. Castelin, personal communication) and one, the Huntsman limestone stromatolites of Macgregor³², some 60 km north of Bulawayo.

Table 2 summarises the western and eastern successions of the Upper Greenstones. Figure 2 shows their most likely distribution in the craton as a whole^{1,33,34}.

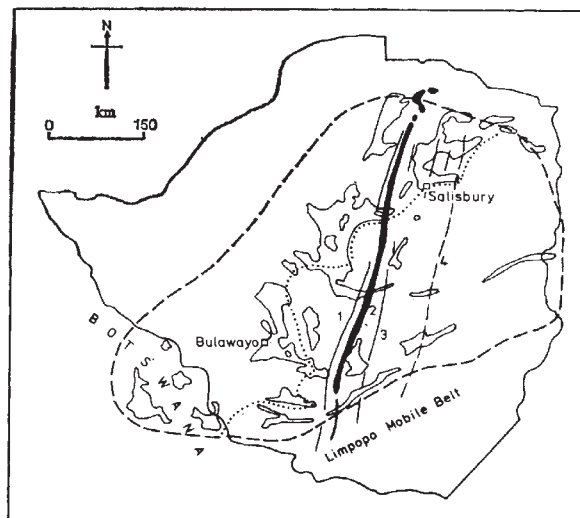
The Shamvaian Group unconformably overlies the Upper Greenstones in the west, east and south-east of the craton¹. The characteristic main rock types are arkoses and sub-greywacke types, rapidly deposited in shallow water and derived from a mixed granitic-volcanic terrain.

Late granites and mafic-ultramafic intrusions

Two suites of late granitic rocks can be recognised post-dating the Upper Greenstones and the Shamvaian Group (Fig. 3). The Sesombi Suite¹ of tonalites and associated granodiorites occurs in the west parallel to the north-north-east trending line of the upper units of the western succession of the Upper Greenstones. The ~ 2,700 Myr ages and the low initial ⁸⁷Sr/⁸⁶Sr ratios of the Sesombi and Somabula tonalites, the two members of the suite so far dated, are indistinguishable from those obtained from the (Maliyami) calc-alkaline volcanic rocks themselves (Table 1). The two intrusions, and by analogy the whole suite, can be regarded as late plutonic expressions of the same magmatic episode which produced the andesites and dacites of the calc-alkaline assemblage. The low initial ratios preclude sources for these volcanics rocks and tonalites in remelting of earlier granite-gneiss terrain^{4,6,35}.

The later ~ 2,600-Myr-old Chilimanzi Suite of intrusions (Table 1) constitute the last major pre-Great 'Dyke' granitic event

Fig. 2 The main greenstone belts of the Rhodesian Archean craton with the Great 'Dyke' and its associated intrusions and major fractures. (1) Umvimeela Dyke; (2) Great 'Dyke'; (3) East Dyke; (4) Popoteke Fault. The dotted line indicates the division between the western and eastern successions of the Upper Greenstones.



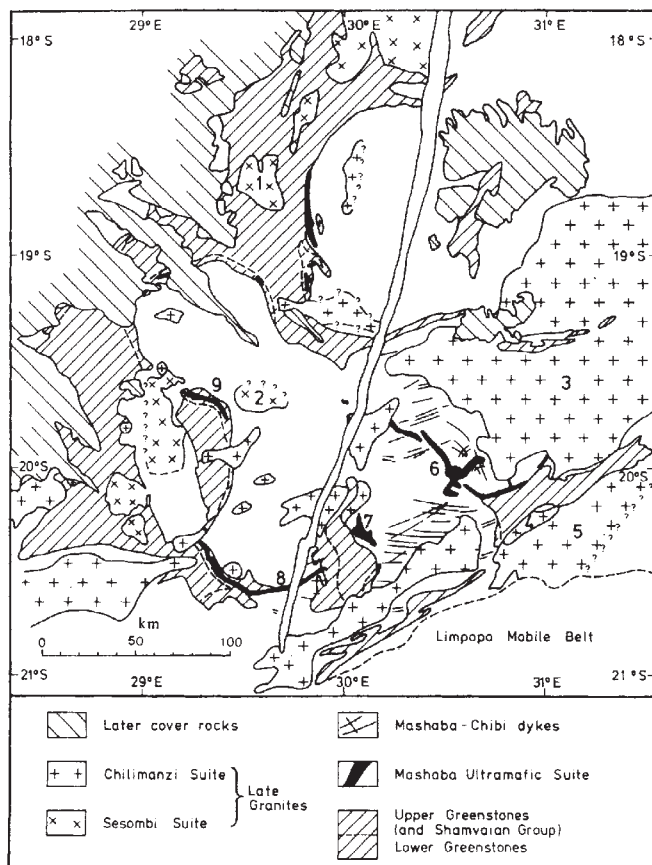


Fig. 3 Distribution of late granites, Mashaba-Chibi dykes and Mashaba Ultramafic Suite in relation to the main greenstone belts in central area of Rhodesia Archaean craton. 1, Sesombi tonalite; 2, Somabula tonalite; 3, Chilimanzi batholith; 4, Victoria Porphyritic granite; 5, Zimbabwe batholith; 6, Mashaba Igneous Complex; 7, Shabani intrusion; 8, Gurumba Tumba-Filabusi intrusion; 9, Shangani intrusion.

in the craton. Petrologically the rocks are dominantly adamellites with local development of granodiorite. The intrusions occur across the craton, but are particularly extensive in the southern and eastern parts. In the southern part of the craton in particular, the intrusions show an east-north-east trend parallel to the Limpopo mobile belt. The strontium isotope data available are compatible with an origin for the Chilimanzi Suite in the remelting of earlier sialic crust^{1,4,36}.

The Mashaba Ultramafic Suite of layered intrusions, together with certain dyke swarms collectively termed the Mashaba-Chibi dykes (Fig. 3), are considered to be broadly contemporaneous with the thick tholeiitic (and high-magnesium) pile which forms the lower part of the Upper Greenstones. None of the intrusions or dykes has yet been dated, but available field evidence places them clearly between the Chilimanzi Suite (~ 2,600 Myr) and the ~ 2,900-Myr-old granitic terrain; they are apparently confined to rocks which pre-date the Upper Greenstones.

The best-documented of the layered intrusions are those of Mashaba and Shabani³⁷⁻⁴⁰, which cut the crustal segment containing the major development of ~ 3,500-Myr-old rocks. The major development of dyke swarms is also within this segment. Some constitute a modified ring and radial pattern and are late phases of the Mashaba Igneous Complex^{37,38}; others are later than the radial swarms and strike east-north-east. The dykes and the Mashaba Igneous Complex also intrude the Mashaba tonalite in this crustal segment.

To the west, dykes and the Gurumba-Tumba Filabusi intrusion cut the Lower Greenstones (Hokonui formation) of the Belingwe belt as well as the Chingezi tonalite and gneisses^{19,23}. North-west of Shangani an undated stock of Sesombi-Suite type truncates the Shangani intrusion¹.

Discussion

The Rhodesian Archaean craton contains granitic rocks intruded during at least three separate events ~ 3,500 Myr, ~ 2,900 Myr and 2,700–2,600 Myr. The ~ 3,500-Myr- and the 2,700–2,600-Myr-old events are both associated with greenstone volcanicity and sedimentation. The ~ 2,900-Myr event may well be associated with the Lower Greenstones, if the Chingezi gneisses and Mashaba-type tonalites represent plutonism complementary to the Lower Greenstones felsic volcanicity.

The status of the Gwenoro and Umniati River gneisses is still obscure. They give ages younger than 2,900 Myr ($2,780 \pm 60$ and $2,760 \pm 80$ Myr, Table 1). The Mashaba Ultramafic Suite and the Mashaba-Chibi dykes serve largely to delineate the pre-Upper Greenstones basement. In neither of these two areas of gneisses, however, have representatives of these intrusions or dykes been recognised. Thus these gneisses could be related to an early phase of the ~ 2,700 Myr-old Sesombi Suite. Alternatively they could equally represent a late stage of tonalitic plutonism complementary to the Lower Greenstones, since Stowe's¹¹⁻¹³ mapping supports a pre-Upper Greenstones age for the Gwenoro area gneisses.

The stratigraphic correlations proposed by J.F.W.¹ emphasise that the Upper Greenstones are the remains of a widespread cover sequence which was subsequently deformed to give, very largely, the present configuration of the main greenstone belts. There is now clear evidence of an extensive basement to this Upper Greenstones cover. This basement consisted of the ~ 3,500-Myr-old granite-greenstone terrain, gneisses and tonalites intruded at ~ 2,900 Myr, the Lower Greenstones, and possibly gneisses dated at ~ 2,800 Myr. Remnants can be seen as far east as the (pre-Mushandike granite) gneisses near Fort Victoria, to the Lower Greenstones on the west side of the greenstone belt north of Bulawayo, and east of Que Que (Fig. 1).

The two differing volcanic provinces recorded in the Upper Greenstones would, on any modern analogy, correspond to the constructive and destructive phases of a plate cycle. Perhaps some of the closest modern analogues to the basal sedimentary and volcanic formations are in rifted continental areas, such as Svartenhuk in west Greenland and the Karoo Super Group in southern Africa. Perioditic lavas are apparently restricted to the Archaean, but these younger rifted areas do contain high-magnesium basalts at the base of thick, near sediment-free, piles of basaltic volcanics^{41,42}. What is not clear, however, is whether a tectonic regime responsible for rifting in the Archaean also produced major ocean basins elsewhere, as happened in these Phanerozoic examples.

The present areal distribution of volcanic rocks in the younger parts of the Upper Greenstones, on the other hand, is reminiscent of modern destructive continental margins (Fig. 1); the true western extent of the calc-alkaline suite is obscured by younger rocks, but on present exposures there is a characteristic linear distribution. It is debatable, however, whether this necessitates an origin in a regime analogous to present-day plate tectonics involving subduction. Hawkesworth and O'Nions³⁵, for instance, point out that, in marked contrast to modern plate margin examples, the tholeiites and calc-alkaline rocks of the Upper Greenstones have trace element characteristics which suggest derivation from chemically similar sources. They favour rifting within a continental block for both rock suites which is compatible also with the wide extent of the pre-Upper Greenstones basement as now established. Nevertheless it is intriguing that if a plate tectonic model is adopted, the Upper Greenstones could record both the creation and subsequent destruction of an Archaean ocean basin.

In considering rifting models it is interesting to observe the re-occurrence through time of a near north-northeasterly trend in the Rhodesian Archaean. This is the tectonic trend of the Upper Greenstones, in particular the calc-alkaline suite; it reflects the fundamental grain of the ancient ~ 3,500-Myr-old rocks; and is seen again in the ~ 2,900-Myr-old Mashaba tonalite. More significantly perhaps it is the trend of the Great 'Dyke' fracture system¹. This includes the line of layered mafic-ultramafic

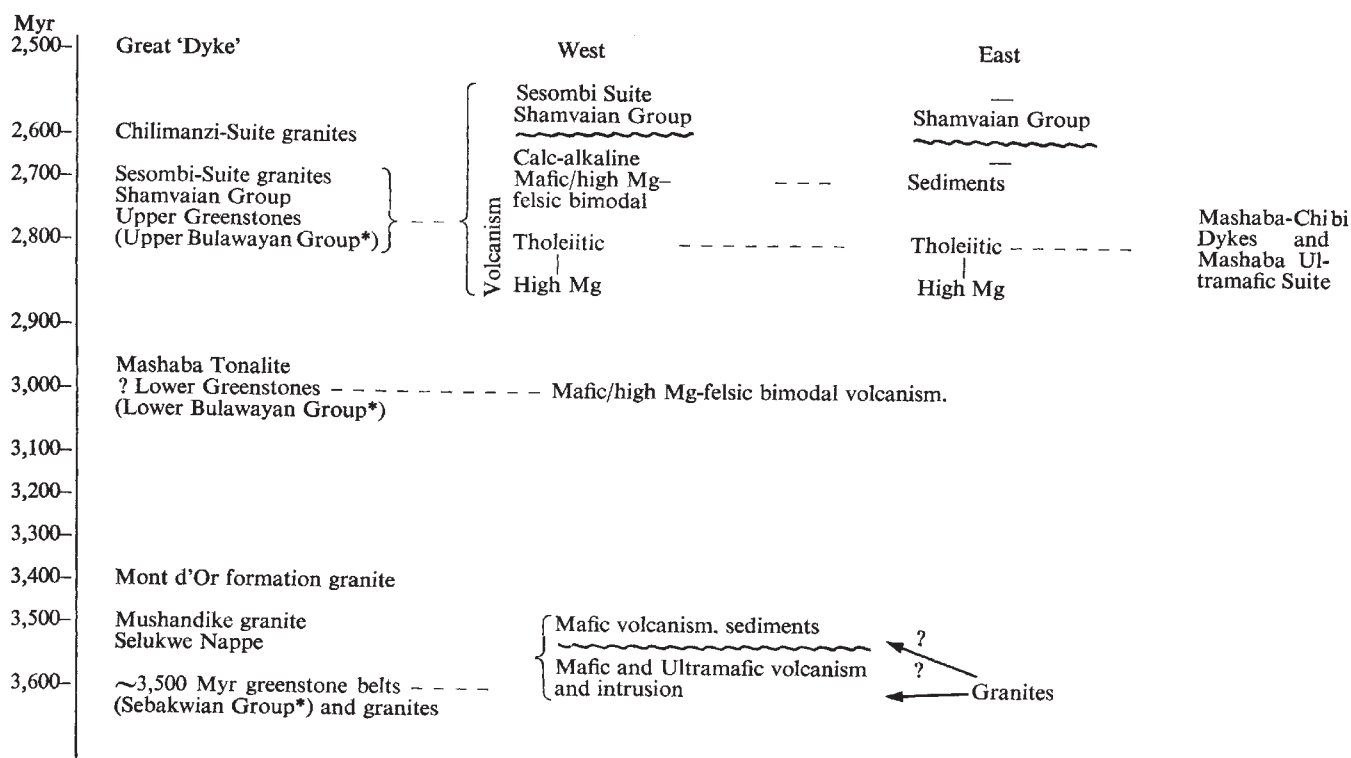


Fig. 4 Summary of major events in the Rhodesian Archaean craton¹. *, Nearest equivalent in Rhodesia Geological Survey terminology.

intrusions which constitute the Great 'Dyke' itself⁴³; the Umvimeela and East (true) dykes; and a number of pre-East dyke sinistral wrench faults of which the Popoteke fault is the longest (Fig. 2).

Pretorius (in ref. 44) has suggested that the Great 'Dyke' formed in a fracture system which developed in the regional block of the Rhodesian craton during movements in adjoining mobile belts. Katz⁴⁵ has proposed that the Limpopo mobile belt is a transform zone which, during its initiation, developed the greenstone belts on the Rhodesian and Kaapvaal cratons in a rift-ridge system at high angles to it, with subsequent reactivation of this transform zone to produce the rift of the Great 'Dyke'. Others⁴⁶ have looked to large intracratonic block movements to explain deformation in the Limpopo belt and within the Rhodesian craton.

Accepting a rifting model for the initiation of the Upper Greenstones, and bearing in mind that these ~ 2,700-Myr-old rocks have not been recognised south of the Limpopo belt in the Kaapvaal craton, it seems reasonable also to accept that major intracratonic block movements across what is now the Limpopo mobile belt (and perhaps even also involving a sub-parallel belt to the north¹) were accompanied by craton-wide fracturing in an approximately north-north-east trend confined to the Rhodesian block, which was effectively the pre-2,700 Myr craton. If this is so then intracratonic block movements, whatever might have been their fundamental causes, have been operative with different degrees of intensity from about 2,700 Myr to 2,500 Myr ago. If the Lower Greenstones had an origin in similar earlier movements, and some of the Limpopo gneisses are ~ 2,900-Myr-old³⁶, then the period is possibly 3,000 Myr to 2,500 Myr. By ~ 2,500 Myr the Rhodesian block, finally stabilised by the widespread Chilimanzi Suite of late granites, fractured sufficiently to produce only the Great 'Dyke' pattern; and thus, in the Great 'Dyke' itself, the magma of which was some kind of magnesium-rich tholeiite^{48,49}, we are perhaps seeing a last abortive attempt at a Rhodesian Archaean greenstone belt^{1,50}.

Figure 4 summarises the major events in the developments of the Rhodesian Archaean craton.

This paper forms part of the Archaean Crustal Study of the Geology Department of the University of Rhodesia, in cooper-

ation with the Universities of Leeds and Oxford. We thank the Director, Rhodesia Geological Survey, for the use of the unpublished material of A. Martin, D. Edwards, J. D. G. Muirhead, I. D. M. Robertson and M. Hickman, and S. Moorbath, R. K. O'Nions and J. Sutton for their help and interest.

Received 31 August; accepted 4 November 1977.

1. Wilson, J. F. *Trans. geol. Soc. S. Afr.* (in the press).
2. Wiles, J. W. *Provisional geological map of Rhodesia 1: 1 000 000* (Rhodesia Geological Survey, Salisbury, 1971).
3. Anhaeusser, C. R. *Minerals Sci. Engng.* **8**, 46 (1976).
4. Hawkesworth, C. J. & Bickle, M. J. *20th A. Rep. res. Inst. Afr. Geol. Univ. Leeds* (1976).
5. Hawkesworth, C. J., Bickle, M. J., Hickman, M. H., Wilson, J. F. & Orpen, J. L. (in preparation).
6. Hawkesworth, C. J., Moorbath, S., O'Nions, R. K. & Wilson, J. F. *Earth Planet. Sci. Lett.* **25**, 251 (1975).
7. Jahn, B. M. & Condie, K. C. *Contrib. mineral. Petrol.* **57**, 317 (1976).
8. Moorbath, S., Wilson, J. F., Goodwin, R. & Humm, M. *Precambrian Res.* **5**, 229 (1977).
9. Hickman, M. H. *Nature* **251**, 259 (1974).
10. Hickman, M. H. *18th A. Rep. res. Inst. Afr. Geol. Univ. Leeds* (1974).
11. Stowe, C. W. *Bull. geol. Surv. Rhod.* **59**, (1968).
12. Stowe, C. W. *Trans. geol. Soc. S. Afr.* **71**, 53 (1968).
13. Stowe, C. W. *J. geol. Soc.* **130**, 411 (1974).
14. Cotterill, P. *Econ. geol. Monogr.* **4**, 154 (1966).
15. Cotterill, P. thesis, Univ. Rhodesia (1976).
16. Moorbath, S., Wilson, J. F. & Cotterill, P. *Nature* **264**, 536 (1976).
17. Wilson, J. F. *Phil. Trans. R. Soc. A* **273**, 389 (1973).
18. Wilson, J. F. & Harrison, N. M. *Spec. Publ. geol. Soc. S. Afr.* **3**, 69.
19. Orpen, J. L. *20th A. Rep. Inst. Afr. Geol. Univ. Leeds* (1976).
20. Bickle, M. J., Martin, A. & Nisbet, E. G. *Earth planet Sci. Lett.* **27**, 133 (1975).
21. Martin, A. *Ann. geol. Surv. Rhod.* **1**, 35 (1975).
22. Bickle, M. J. & Nisbet, E. G. *Geology* February 68 (1977).
23. Martin, A., Bickle, M. J., Nisbet, E. G. & Orpen, J. L. (in preparation).
24. Nisbet, E. G., Bickle, M. J. & Martin, A. *J. Petrol.* (in the press).
25. Worst, B. G. *Bull. geol. Surv. Sth. Rhod.* **43**, (1956).
26. Macgregor, A. M. *Bull. geol. Surv. Sth. Rhod.* **31**, (1937).
27. Macgregor, A. M. *Bull. geol. Surv. Sth. Rhod.* **38**, (1947).
28. Macgregor, A. M. *Proc. geol. Soc. S. Afr.* **54**, (1951).
29. Harrison, N. M. *Bull. geol. Surv. Rhod.* **67**, (1970).
30. Condie, K. C. & Harrison, N. M. *Precambrian Res.* **3**, 253 (1976).
31. Bliss, N. W. *Bull. geol. Surv. Rhod.* **64**, (1970).
32. Macgregor, A. M. *Trans. geol. Soc. S. Afr.* **43**, 9 (1941).
33. Wiles, J. W. *Bull. geol. Surv. Rhod.* **63**, (1972).
34. Key, R. M., Litherland, M. & Hepworth, J. V. *Precambrian Res.* **3**, 375 (1976).
35. Hawkesworth, C. J. & O'Nions, R. K. *J. Petrol.* **18**, 487 (1977).
36. Hickman, M. H. thesis, Univ. Leeds (1976).
37. Wilson, J. F. *Bull. geol. Surv. Rhod.* **62**, (1968).
38. Wilson, J. F. *Trans. geol. Soc. S. Afr.* **71**, Annex 175 (1968).
39. Laubscher, D. H. in *The Geology of Some Ore Deposits of Southern Africa* (ed. Haughton, S. H.) 593 (Geological Society of South Africa, 1964).
40. Laubscher, D. H. *Trans. geol. Soc. S. Afr.* **71**, Annex 195 (1968).
41. Drever, H. I. & Johnston, R. *Trans. R. Soc. Edin.* **63**, 289-314 (1957).
42. Cox, K. G. *Phil. Trans. R. Soc.* (in the press).
43. Worst, B. G. *Bull. geol. Surv. Sth. Rhod.* **47**, (1960).
44. Sutton, J. *Spec. Publ. geol. Aust.* **3**, 1 (1971).
45. Katz, M. B. in *The Early History of the Earth* (ed. Windley, B. F.) 147 (Wiley, New York, 1976).
46. Coward, M. P., James, P. R. & Wright, L. *Bull. geol. Soc. Am.* **67**, 601 (1976).
47. Hamilton, P. J. *J. Petrol.* **18**, 24 (1977).
48. Robertson, I. D. M. & van Breemen, O. *Spec. Publ. geol. Soc. S. Afr.* **1**, 621 (1970).
49. Wilson, A. H. thesis, Univ. Rhodesia (1977).
50. Coward, M. P., Lintern, B. C. & Wright, L. I. in *The Early History of the Earth* (ed. Windley, B. F.) 33 (Wiley, New York, 1976).

Direct comparison of theoretical and experimental melting profiles for RF II Φ X174 DNA

Y. L. Lyubchenko, A. V. Vologodskii & M. D. Frank-Kamenetskii

Department of Biology, I. V. Kurchatov Institute of Atomic Energy, Moscow, USSR

The determination by Sanger et al. of the complete nucleotide sequence for Φ X174 DNA has made it possible for the first time to compare directly theoretical and experimental DNA melting profiles. The comparison shows that the theory predicts the observed shape of the differential melting curve surprisingly well. Calculation of the denaturation maps allows the peaks on the curve to be correlated with co-operative melting out of concrete regions on the sequence of nucleotides.

HELIX-coil transition or intramolecular melting of the DNA double helix has been studied extensively both experimentally and theoretically for many years (see refs 1-4 for reviews). Until recently the DNA melting curve was thought to reflect rather rough characteristics of base-pair distribution and by the use of it a DNA sequence could be placed only into one of the two categories—quasirandom or block-quasirandom (see refs 1, 2). Recently, however, the viral DNA melting curves have been shown to exhibit a clear-cut fine structure which reflects specific DNA sequences to a much greater extent⁴⁻¹⁰. This fine structure manifests itself most clearly when a derivative of the degree of denaturation with respect to temperature, or differential melting curves, are plotted. Such a plot for a sufficiently short genome consists of a series of sharp peaks with half-width of about 0.3° (see refs 5-10).

There has been considerable progress in the construction of a theory of helix-coil transition in DNA. The theory predicts correctly the main overall characteristics of the process—the melting range width, the number of bases in alternating helical and coiled sections within the melting zone and others. Moreover, the theory has succeeded in predicting the effect of fine structure of the differential melting profiles^{4,5,11}. Sequences in these calculations were simulated with the aid of a generator of random digits because of the lack of data on nucleotide sequence for any DNA: this prevented the direct comparison between the theory and experiment. The situation has changed now that Sanger et al.¹² have published the complete nucleotide sequence for Φ X174 DNA. In the present work we are using these data to compare directly, for the first time, theoretical and experimental differential melting profiles.

Recent advances in theory

A model that is used in the theory of helix-coil transition is rather simple (see refs 1-4 for reviews). Each base pair is assumed to belong to one of the two states—'helical' or 'coil-like'. Opening of a base pair at the end of a helix needs free energy ΔF_{AT} for an AT pair and ΔF_{GC} for a GC pair. Formation of a new unwound region inside the helix needs an additional energy, F_s , that may be considered as the energy of base-pair stacking. In addition to these ordinary features of the well known Ising model, the DNA model includes so-called loop-weighting factors. These statistical weights appear because unwound regions in DNA have the shape of a closed polymer chain.

Comparison of theoretical and experimental melting profiles

Closed circular double-stranded replicative form (RFI) of Φ X174 DNA is not of interest for our purposes because both theoretical

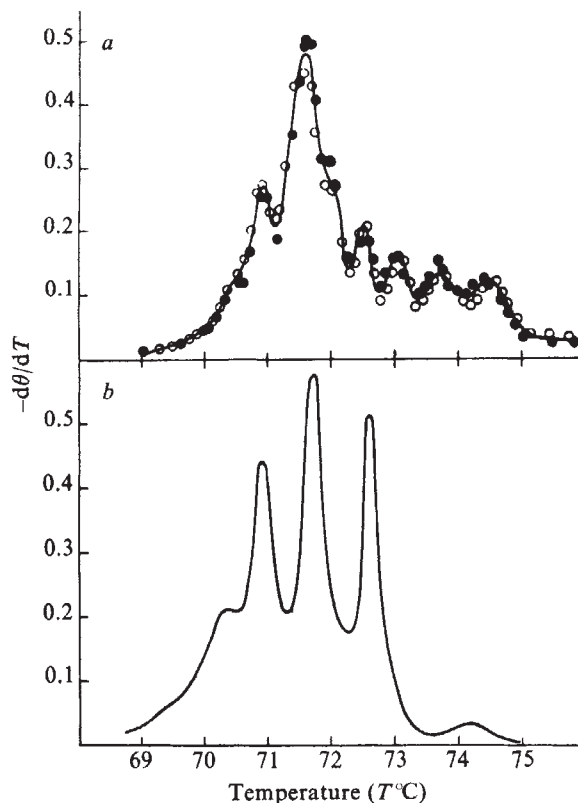


Fig. 1 Theoretical (b) and experimental (a) differential melting profiles for RF II Φ X174 DNA in 0.1SSC. The theoretical curve was calculated for the sequence of nucleotides determined by Sanger et al.¹² using the theory of helix-coil transition in DNA using the standard set of parameters^{1,2,4,5,18}: $\sigma = 5 \times 10^{-5}$; $\alpha = 1.5$; $U_{AT} = 8 \text{ kcal/mol}^{-1}$, $T_{AT} = 52.5^\circ\text{C}$; $T_{GC} - T_{AT} = 42.4^\circ\text{C}$. Replicative form (RFI) of Φ X174 DNA was obtained from the cells of *Escherichia coli* C infected by phage Φ X174, the multiplicity of infection about 5-10, with subsequent treatment by chloramphenicol. The phage Φ X174 of wild type and *E. coli* C cells were gifts of Drs E. Stukacheva and L. Patrushev. The chromosomal DNA was removed from the suspended cells by centrifugation (40,000g, 1 h)²¹ with subsequent extraction by 50% isopropanol. The separation of RFI DNA from the mixture of RFI and the linear DNA molecules was done by the centrifugation in density gradient of CsCl with addition of ethidium bromide (200 $\mu\text{g/ml}^{-1}$)²². Ethidium bromide was removed by threefold isopropanol extraction and subsequent dialysis of DNA against SSC buffer (0.15M NaCl, 0.015M Na citrate), pH7. The sample contained about 90% of RFI DNA and 10% of RFII DNA. Conversion of RFI DNA into RFII DNA was done using S1 endonuclease¹⁹ in 0.2M CH_3COONa , pH4.6 in the presence of 10^{-4}M ZnSO_4 at 37°C ; sufficient enzyme was added to half of the denatured DNA (5 $\mu\text{g/ml}^{-1}$) in 10 min. The reaction was followed by the disappearance of RFI band in the course of Agarose gel electrophoresis. Incubation for 30 min led to the complete disappearance of the RFI DNA. At the same time about 10% of the molecules were converted into linear form. Melting of DNA was followed by the disappearance of RFI band in the course of Agarose gel electrophoresis. Incubation for 30 min led to the complete disappearance of the RFI band. Differential melting profiles were obtained by graphical differentiation and automatically²³. The profiles obtained by the different methods and in many independent measurements were very similar. The filled and open circles on the experimental curve (a) present two examples of such measurements. The long plateau in the region of high temperatures represents contamination, probably due to host DNA.

predictions and experimental observations led to the conclusion that such DNA molecules should have broad but structureless differential melting profiles⁵. So theoretical calculations and experimental investigations should be conducted with RF II DNA having a single-stranded break (nick).

We have converted RFI DNA into RF II using S1 endonuclease. This enzyme was shown by Kato *et al.*¹⁹ to introduce the nicks into RF I Φ X174 DNA by a random way. The theoretical profile was obtained by averaging over ten calculated curves differing one from another by a position of the nick. The shape of the theoretical differential melting profile depends on the position of the nick but this dependence is not strong—for the peaks are shifted by not more than by 0.1–0.2 °C.

Theoretical and experimental differential melting profiles thus obtained are given in Fig. 1. In general, the curves are remarkably alike. Indeed, positions of peaks at 70.4, 70.9, 71.7 and 72.6 °C coincide for theoretical and experimental profiles with as high precision as of about 0.1 °C. But the amplitudes of the peak at 72.6 °C are very different for both curves and, beginning with this peak, the curves differ from one another considerably. We shall discuss the possible causes of these discrepancies (see below).

Correlation of theoretical denaturation maps with melting profile

In addition to the melting profile, the theory permits the calculation of a denaturation map, that is, the dependence of the probability of a base pair to be open on its position within a nucleotide sequence.

Figure 2 shows theoretical denaturation maps calculated for RF II Φ X174 DNA in different temperatures. Values of the temperature are indicated to the right. These maps being compared with the differential melting profile presented in Fig. 1 enable us to correlate the peaks on the profile with co-operative melting of concrete regions on the nucleotide sequence published by Sanger *et al.*¹².

Until recently a rigorous and effective algorithm for calculating a melting curve for DNA with an arbitrary sequence of base pairs has been lacking. Such an algorithm has now been formulated as the result of the following sequence of events.

Lacombe and Simhar¹³ developed the unique probability formalism allowing the calculation of equilibrium characteristics for a linear heterogeneous system. This formalism was different from, but equivalent to the ordinary formalism based on the partition function method. Poland¹⁴ showed that this formalism made it possible to treat rigorously the DNA model. But, the method of Poland, and the previously proposed method of Wartell and Montroll³ based on ordinary formalism, were not sufficiently effective—the computer time required was proportional to N^2 where N is the number of base pairs in the chain. An approximate method¹⁵ was proposed as early as 1969 for which the computer time is proportional to N . This method has been used extensively in this laboratory^{1,2,5}. Recently Lukashin *et al.*¹⁶ have compared the results of calculations by this method with the results obtained by the exact methods^{3,14} and concluded that the approximation¹⁵ gives results which are in accord with results of exact methods. These findings stimulated Fixman and Freire¹⁷ to introduce an idea that formed the basis of the method¹⁵ into formalism^{13,14}.

Fixman and Freire¹⁷ succeeded in formulating an exact algorithm for which computer time is proportional to the number of base pairs in DNA. So now, at last, theoreticians have an exact and highly effective method¹⁷ for treating the DNA model. This method is used here to calculate differential melting profiles and denaturation maps for RF II Φ X174 DNA.

Parameters of the theory

The theory of helix-coil transition makes it possible to compute a melting curve and a denaturation map for DNA if a sequence of base pairs and values of parameters of the theory are given. The theoretical parameters are as follows (see refs 1–4): the melting enthalpy for an AT pair, U_{AT} ; the melting temperature of

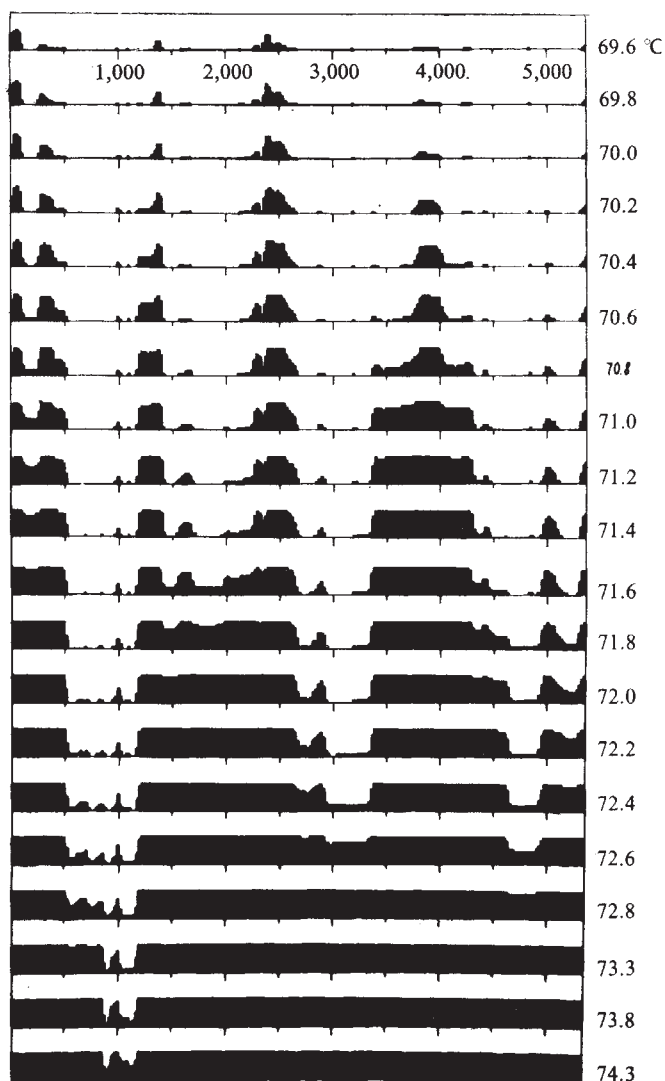


Fig. 2 Theoretical denaturation maps for RF II Φ X174 DNA calculated with the same set of parameters as indicated in the legend for Fig. 1. As for the melting profile, each map was obtained by averaging over ten maps differing from one another by the position of the nick. The sequence is numbered as in Fig. 1 of ref. 12. An ordinate presents the probability of a base pair to be open averaged over 25 subsequent base pairs.

poly AT; T_{AT} ; the difference $T_{GC} - T_{AT}$; the co-operativity factor, σ ($\sigma = \exp(-F_s/RT)$); and loop-weighting exponent α .

The value U_{AT} has been determined from calorimetric and other data and it is equal to 8.0 kcal/mol (see, for example, ref. 4). The values of T_{AT} and $T_{GC} - T_{AT}$ depend on ionic conditions and are determined from a series of experiments with DNA having different GC content. For conditions used in the present work (0.1SSC), $T_{AT} = 52.5$ °C and $T_{GC} - T_{AT} = 42.4$ °C, see ref. 18.

The value of σ is chosen as a rule to be equal to 5×10^{-5} and $\alpha = 1.5$ (see refs 1,2,4,5). A considerable ambiguity exists in this choice of values—the value of $\alpha = 1.8$ is more justifiable from the theoretical point of view than the value of 1.5 and that of σ is not known more accurately than to within an order of magnitude. Fortunately, as appropriate calculations have shown, the differential melting profile, at least in our case, is practically unaffected by the change of these two parameters within the limits of their possible variations. In general, our calculations show that the shape of differential melting curve is highly sensitive to the sequence of base pairs and to the values of parameters T_{AT} and $T_{GC} - T_{AT}$ but insensitive to the particular values of the rest of parameters.

The regions melting at the lowest temperatures are too short to give distinct peaks on the differential melting profile. Nevertheless the shoulder at 69.3 °C may be recognised on the theoretical profile originating from the melting out of the region, which is situated, as one may conclude from Fig. 2, between bases with order numbers 25 and 100. Inspection of the corresponding piece of the nucleotide sequence¹² shows that this region is highly enriched (by about 20%) by AT pairs as compared with the sequence as a whole.

The first distinct peak at 70.4 °C appears mainly due to a co-operative melting of the region situated between 3.80 and 4.00 kilobases. Then the two neighbouring regions situated at 3.35–3.80 and 4.00–4.30 kilobases are melting out simultaneously giving the large peak at 70.9 °C. The main peak at 71.7 °C originates from a co-operative melting of the very large region containing about 850 base pairs situated between 1.40 and 2.25 kilobases. Simultaneously the region between 4.30 and 4.65 kilobases melts. The melting of this region that may explain, at least in part, the peak at 72.0 °C on the experimental profile. Indeed, such a peak, but with much lower magnitude than the experimental one, appeared on theoretical profiles for some positions of the nick.

The large peak at 72.6 °C on the theoretical profile corresponds to a simultaneous melting of two independent regions at 2.90–3.35 and 4.65–4.95 kilobases. Finally the last small peak on the theoretical profile originates from the melting of the region positioned between 1.00 and 1.15 kilobases. The small area which remains completely helical up to the highest temperature is situated, as it can be seen from Fig. 2, at 0.87–0.90 kilobases. Inspection of the nucleotide sequence¹² in this site shows that starting from 872-th nucleotide the sequence of 35 bases highly enriched by GC pairs follows. It contains only 8 AT pairs and it has higher GC content, compared with the mean value, by more than 25%. This GC cluster remains completely helical up to about 73.5 °C.

It should be remembered in this connection that the theory of helix-coil transition in DNA used here is valid under the assumption that complete dissociation of complementary strands does not take place. This assumption is valid until a single helical region exists, that is, up to about 73.5 °C, for our particular case, and it fails for higher temperatures. So, at best, accordance between the theory and experimental results may be expected only for the temperatures lower than 73.5 °C. We can conclude that the only serious difference between the theoretical and experimental melting profiles is the too large a magnitude of the peak at 72.6 °C on the theoretical curve. Indeed the melting of the region situated between 0.50 and 0.85 kilobases (see Fig. 2) should give a peak at 72.8 °C, which may be considered as corresponding to the peak at 73.0 °C on the experimental curve. On the theoretical curve (see Fig. 1) it is masked by a powerful peak at 72.6 °C.

It should be emphasised that in general it is absolutely impossible to predict, just by inspecting a sequence of nucleotides, what particular region would melt out co-operatively to result in a peak on the differential melting profile. Only for two extreme regions, highly enriched by AT pairs and highly enriched by GC pairs, does the simple explanation of their behaviour, given above, exist, but these sequences are too short to give a distinct peak on the profile. For the rest of the sequence it is only the theory of helix-coil transition in DNA that makes it possible to indicate the regions which are melting out co-operatively and to correlate them with peaks on the differential melting curve.

Note also that the melting temperature of a region depends not only on its GC content but also on its length and boundary conditions. For example, the region at 0.50–0.85 kilobases contains about 47% GC, the region at 4.65–4.95 kilobases contains about 50% GC but the former melts at a slightly higher temperature than the latter in contrast to the expectation based only on their GC content. The causes are clear—the region at 0.50–0.85 kilobases had a melted area only to its left before melting, but the region 4.65–4.95 kilobases had the melted areas from both sides. Thus considerable errors may follow from an attempt to correlate

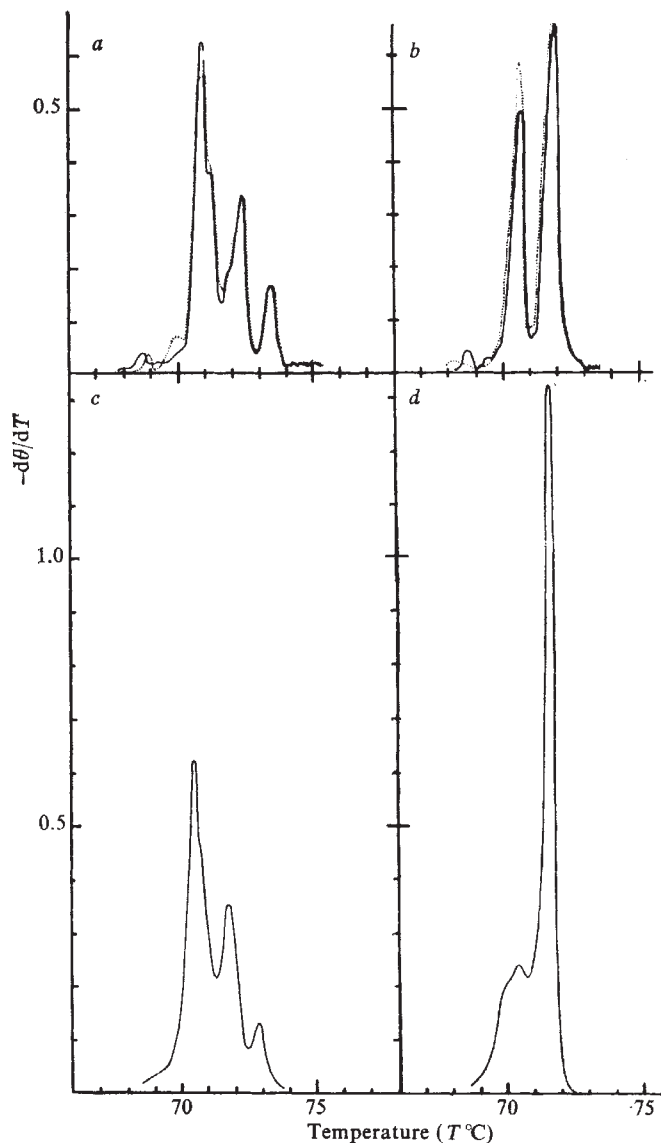


Fig. 3 Theoretical (c, d) and experimental (a, b) differential melting profiles for restriction fragments Y_1 (a, c) and Y_2 (b, d) of $\Phi X174$ DNA in a 0.1 SSC aqueous solution. Experimental profiles are taken from Wada *et al.*²⁷. Solid and dotted curves correspond to different experiments and demonstrate reproducibility. Theoretical curves were calculated using the parameters listed in Fig. 1 legend.

directly the position of a peak on a differential melting curve with GC content of a corresponding region.

Possible causes of discrepancy

The above discussion shows that the only important discrepancy between the theoretical and experimental melting profiles is the difference of magnitudes of the peak at 72.6 °C for both curves. Special calculations excluded an ambiguity in the values of parameters σ and α as a possible cause of this difference. Two causes that may be responsible for it are discussed below.

First the real sequence of nucleotides in the DNA studied by us experimentally may differ from the sequence published by Sanger *et al.*¹². The phages used by Sanger *et al.*¹² and by us are of different origin but they probably differ from one another by only a few point mutations. Possibly of more importance, however, as Sanger *et al.*¹² emphasised, occasional errors might be introduced into the sequence during the laborious experiments involved in the determination. The regions where such errors may occur are shown in Fig. 1 of their paper¹². In the four positions the bases are not indicated at all. We have filled these gaps in the following way: 1,421, A; 1,422, A; 1,423, C and 1,657, C. These few uncertainties cannot affect the results presented above. But our results may be influenced by the fact that a very long region exists starting at 2.95 and extending to 3.68 kilobases where one mistake

per 50 residues, and in some places even greater level of uncertainty, is possible. Indeed, one of the regions contributing to the peak at 72.6 °C on the theoretical curve in Fig. 1, the region 2.90–3.35 kilobases, is situated almost completely within this most uncertain area. Note that most of the theoretical profile presented in Fig. 1 cannot be changed because it is formed by the melting of sequences fully confirmed.

Second, the effect of heterogeneity of stacking interactions between base pairs neglected in the theory, may influence the results. This effect is large for RNA helices according to available data²⁴. But, in the case of DNA this effect is shown to be very small by direct experimentation with DNA²⁵ as well as by the theoretical analysis²⁶ of experimental data on melting of synthetic polynucleotides with known sequences. The above results show the theory to be successful in precisely predicting positions of a series of the peaks on the differential melting curve for Φ X174 DNA; this constitutes evidence for a neglected role of the effect of heterogeneity of stacking interaction. Nevertheless, we cannot yet exclude completely a possibility that in some special situations this effect may give an appreciable contribution.

It is difficult to draw an unambiguous conclusion from this work because of the possible errors in the published sequence. Nevertheless the results obtained strongly suggest that the theory can efficiently predict the shape of a differential melting curve in detail. The model which forms the basis of the theory is thus appropriate to real DNA.

We thank Dr A. V. Lukashin for help, Drs Y. S. Lazurkin and V. I. Ivanov for discussions, Mrs V. N. Syomushkina for technical assistance and Dr M. Fixman for sending a preprint of his paper.

After completion of this study we became acquainted with the very interesting data of Wada *et al.*²⁷. These authors measured

the differential melting curves for two restriction fragments of Φ X174 DNA: Y¹ consisting of 2,745 base pairs (position 3,360–729) and Y² consisting of 1,690 base pairs (position 1,104–2,793). We have calculated the theoretical melting profiles for these fragments which are presented in Fig. 3 along with the experimental profiles taken from ref. 27.

Received 22 July; accepted 20 October 1977.

1. Lazurkin, Y. S., Frank-Kamenetskii, M. D. & Trifonov, E. N. *Biopolymers* **9**, 1253–1306 (1970).
2. Vedenov, A. A., Dykhne, A. M. & Frank-Kamenetskii, M. D. *Usp. Fiz. Nauk* **105**, 479–519 (1971); *Sov. Phys.-Usp.* **14**, 715–736 (1972).
3. Wartell, R. M. & Montroll, E. W. *Adv. Chem. Phys.* **22**, 129–203 (1972).
4. Frank-Kamenetskii, M. D. & Lazurkin, Y. S. *A. Rev. Biophys. Bioeng.* **3**, 127–150 (1974).
5. Lyubchenko, Y. L., Frank-Kamenetskii, M. D., Vologodskii, A. V., Lazurkin, Y. S. & Gause, G. G. *Biopolymers* **15**, 1019–1036 (1976).
6. Yabuki, S., Gotoh, O. & Wada, A. *Biochim. biophys. Acta* **395**, 258–273 (1975).
7. Gotoh, O., Husimi, Y., Yabuki, S. & Wada, A. *Biopolymers* **15**, 655–670 (1976).
8. Wada, A., Tachibana, H., Gotoh, O. & Takamami, M. *Nature* **263**, 439–440 (1976).
9. Ansevin, A. T., Vizard, D. L., Brown, B. W. & McConathy, J. *Biopolymers* **15**, 153–174 (1976).
10. Vizard, D. L. & Ansevin, A. T. *Biochemistry* **15**, 741–750 (1976).
11. Frank-Kamenetskii, M. D. & Vologodskii, A. V. *Nature* **269**, 5630 (1977).
12. Sanger, F. *et al.* *Nature* **265**, 687–695 (1977).
13. Lacombe, R. H. & Simha, V. *J. chem. Phys.* **58**, 1043–1053 (1973).
14. Poland, D. *Biopolymers* **13**, 1859–1871 (1974).
15. Frank-Kamenetskii, M. D. & Frank-Kamenetskii, A. D. *Molekulyarnaya Biologiya* **3**, 375–383 (1969); *Molec. Biol.* **3**, 295–301 (1969).
16. Lukashin, A. V., Vologodskii, A. V. & Frank-Kamenetskii, M. D. *Biopolymers* **15**, 1841–1844 (1976).
17. Fixman, M. & Freire, J. J. *Biopolymers* (in the press).
18. Lazurkin, Y. S., Lyubchenko, Y. L., Pavlov, V. M., Frank-Kamenetskii, M. D. & Berestet-skaya, I. V. *Biopolymers* **14**, 1551–1552 (1975).
19. Kato, A. C., Bartok, K., Fraser, M. J. & Denhardt, D. T. *Biochim. biophys. Acta* **308**, 68–78 (1973).
20. Komano, T. & Sinheimer, R. L. *Biochim. biophys. Acta* **155**, 295–298 (1968).
21. Bazaral, M. & Helinski, D. R. *Proc. natn. Acad. Sci. U.S.A.* **62**, 1159–1166 (1969).
22. Radloff, R., Bauer, W. & Vinograd, J. *Proc. natn. Acad. Sci. U.S.A.* **57**, 1514–1521 (1967).
23. Pavlov, V. M. & Lyubchenko, Y. L. *Biopolymers* (in the press).
24. Borer, P. N., Dengler, B., Tinoco, I. Jr & Uhlenbeck, O. C. *J. molec. Biol.* **86**, 843–853 (1974).
25. Voskoboinik, A. D. *et al.* *Molekulyarnaya Biologiya* **9**, 783–790 (1975).
26. Belintsev, B. N., Vologodskii, A. V. & Frank-Kamenetskii, M. D. *Molekulyarnaya Biologiya* **10**, 764–769 (1976).
27. Wada, A., Tachibana, H., Ueno, S., Husimi, Y. & Machida, Y. *Nature* **269**, 352–353 (1977).

Comparison of amino acid sequence of troponin I from different striated muscles

J. M. Wilkinson & R. J. A. Grand

Department of Biochemistry, University of Birmingham, PO Box 363, Birmingham, UK

The sequence of troponin I from fast and slow skeletal and cardiac muscle shows strong homology in the region which binds to actin and is responsible for inhibition of the actomyosin ATPase. More differences are found in the N-terminal region which binds to troponin C.

CONTRACTILE activity in vertebrate striated muscle is regulated by the level of free Ca²⁺ in the sarcoplasm by means of the molecular switch formed by the troponin complex and tropomyosin. Troponin is composed of three components, troponin I, which inhibits the actomyosin ATPase, troponin C, which binds Ca²⁺ and troponin T, which binds to tropomyosin and locates the complex at a repeat of 385 Å along the thin filament^{1,2}. Troponin I will inhibit the actomyosin ATPase on its own but this inhibition is greatly enhanced in the presence of tropomyosin³. The inhibition is neutralised by troponin C and the restoration of full Ca²⁺ sensitivity requires troponin T^{4,5}. Each component of the thin filament can be shown to interact with two others to bring about this regulation and it has been shown that troponin I interacts with both actin⁶ and troponin C⁷, this latter interaction being Ca²⁺ sensitive. The primary structures of all the major thin filament proteins from rabbit fast skeletal muscle are known^{8–12} but while the pattern of regulation in the other major types of vertebrate striated muscle seems similar, the troponin components have been shown to differ in different muscle types^{13–15} and this may well reflect modifications in

regulatory activity. We have determined the amino acid sequence of four troponin I proteins from different sources in order to investigate the molecular basis of the interaction with the other thin filament proteins, troponin C and actin, and also to elucidate the differences between proteins obtained from different types of striated muscle. The sequences of three of the proteins chosen, fast¹¹ and slow¹⁶ skeletal and cardiac¹⁷ troponin I from the rabbit, have been published and we report here the sequence of troponin I from chicken fast skeletal muscle. A detailed description of the determination of this sequence will be published elsewhere.

Primary structure determination

The amino acid sequences of the four proteins are shown in Fig. 1. The four sequences have been aligned to give maximum homology and it will be seen that a number of deletions have to be introduced into each sequence. The numbering of the residues is that for the cardiac protein, which includes two deletions at positions 181 and 182. One correction must be made to the previously published sequence of the rabbit fast muscle protein¹¹. The sequence in the region 139–143 was previously thought to be Leu-Arg-Arg-Arg-Val but has now been shown to be Leu-Arg-Arg-Val. The corresponding sequence in the rabbit cardiac protein is Leu-Arg-Leu-Arg-Val, this was noted previously¹⁷ to be in some doubt but efforts to clarify this sequence were unsuccessful. A blocked N-terminus was found in each protein.

	10	20	30
C.F.			X-Ser
R.F.			Ac-Gly-Asp-Glu-Glu-
R.S.			NH-Pro Val-
R.C.	X-Ala-Asp-Glu-Ser-Arg-Asp-Ala-Ala-Gly-Glu-Ala-Arg-Pro-Ala-Pro-Ala-Val-Arg-Arg-Ser-Asp-Arg-Ala-Tyr-Ala-Thr-Glu-Pro-His-Ala-		
	40	50	60
C.F.	Lys-Arg Ala	Ala Leu Val Ile	Ala-
R.F.	Lys-Arg-Asn-Arg-Ala-Ile-Thr-ALA —	ARG Arg-Gln-His-LEU-LYS-Ser-Val-MET-LEU-Gln-Ile-ALA-Ala-Thr-Glu-Leu-Glu-Lys-Glu-Glu-	
R.S.	Glu Lys-Ser-Lys Ser Lys-Leu —	Leu Ala-Lys Lys-Glu-Cys-Gln-Gln-Glu-His	
R.C.	Ser-Lys-Lys-Lys Ser Ser Lys-Leu-Gln Thr-Leu	Lys-Gln Arg Ala-	
	70	80	90
C.F.	Ala-Ala-Lys Val		Gln Leu
R.F.	Gly-Arg-Arg-Glu-Ala-GLU-LYS-Gln-Asn-Tyr-LEU-Ala-Glu-His-Cys-Pro-Pro-LEU-Ser-Leu-Pro-GLY —	Ser-Met-Ala-Glu-Val-GLN-Glu-	
R.S.	Ala — Val-Arg Arg-Ile Ala Gln-Thr-Arg Leu Leu-Ser-Ala-Leu Asp-		
R.C.	Glu-Glu Arg-Gly Gly-Arg-Ala Ser-Thr-Arg Gln Glu Ala Leu-Gly-Phe Leu Asp		
	100	110	120
C.F.	Lys Ser-Val-Asp Arg Thr Val Leu Thr-Asn		
R.F.	LEU-CYS-Lys-Gln-LEU-HIS-ALA-Lys-Ile-Asp-Ala-Ala-Glu-GLU-GLU-Lys TYR-ASP-Met-GLU-Ile-LYS-Val-Gln-Lys-Ser-Ser-Lys-GLU-Leu		
R.S.	Arg Val-Glu-Val-Val-Asp Arg Ile Ala Cys-Leu-His-Asn-Thr-Arg Ile		
R.C.	Arg Arg-Val Lys-Val-Asp Arg Val Ala Thr Asn-Ile-Thr Ile		
	130	140	150
C.F.	Leu-Ser		
R.F.	Glu-ASP-Met-Asn-Gln-LYS-Leu-Phe-ASP-LEU-ARG-GLY-LYS-PHE-LYS-ARG-PRO-Pro-LEU-ARG —	ARG-VAL-ARG-Met-SER-ALA-ASP-ALA-MET	
R.S.	Lys Leu-Lys-Leu Val-Leu		Val
R.C.	Ala Leu-Thr Ile Thr Leu Ile		
	160	170	180
C.F.	Arg Asn		
R.F.	Leu-Lys-ALA-LEU-LEU-GLY-Ser-Lys-His-LYS-Val-Cys-Met-ASP-LEU-ARG-ALA-Asn-LEU-LYS-Gln-VAL-LYS-LYS-GLU -ASP-THR-GLU-LYS-GLU		
R.S.	Arg Ser Ser		
R.C.	Met-Gln Thr-Arg-Ala Glu-Thr-Leu His		
	190	200	210
C.F.	Lys-Asp-Leu		Ala-Gly
R.F.	— — — Arg-Asp-VAL-GLY-ASP-TRP-ARG-LYS-ASN-Ile-Glu-Glu-Lys-SER-GLY-MET-GLU-GLY-ARG-LYS-LYS-Met-PHE-Glu-Ser —	Glu	
R.S.	Arg-Pro-Val — Glu Val Ala-Met		Asp-Ala-Ala-Lys
R.C.	Asn Glu Asp-Leu-Leu Lys Gly- COOH		
	215		
C.F.	- COOH		
R.F.	Ser- COOH		
R.S.	Pro-Thr-Ser-Gln- COOH		

Fig. 1 Alignment of the amino acid sequences of troponin I from chicken fast muscle (C.F.), rabbit fast muscle (R.F.), rabbit slow muscle (R.S.) and rabbit cardiac muscle (R.C.). The complete sequence of rabbit fast muscle troponin I is presented, gaps in the other sequences indicate identity with this sequence. Residues which are identical in all four sequences are shown in capitals. Solid lines represent deletions. X indicates an unknown blocking group.

In the rabbit fast muscle protein this has been shown to be an acetyl group¹¹ and this is likely to be the case for the other proteins as all other myofibrillar proteins so far studied have been shown to be N-acetylated. One curious feature of the rabbit slow muscle protein is that although some of the molecules have a blocked N-terminus, most have proline with a free imino group and we believe this to be the first example of a myofibrillar protein with a free N-terminus. Heterogeneity was also found at the C-terminal

end of the molecule where a proportion of the molecules lacked the terminal Ser-Gln dipeptide. The origin of these different forms of slow muscle troponin I is not clear, but as the protein was isolated by affinity chromatography after extraction of muscle with 9 M urea it is unlikely that they are an artefact of preparation. Two other myofibrillar proteins, the A1 and the 'phosphorylatable' light chains of myosin^{18,19}, have been shown to have a blocked N-terminal proline, but in both cases blocking seemed to be complete.

Comparison of sequences

The four complete troponin I sequences have been compared by calculating the percentage difference between pairs of proteins (Fig. 2). As the proteins differ slightly in length, extra residues at either the N- or C-terminus were ignored and deletions were treated as an extra amino acid. The proteins were also divided arbitrarily between residues 119 and 120 and the N- and C-terminal parts were compared. The matrices (Fig. 2) show that the fast muscle proteins from both chicken and rabbit are rather closely related and have only half the number of differences found between any two of the three rabbit proteins. In the case of the three rabbit proteins the number of differences between any two of them are approximately the same. When the N- and C-terminal halves are compared this general pattern is still maintained but mutations have accumulated in the N-terminal half at approximately twice the rate observed in the C-terminal half. There is some indication that the N-terminal halves of the rabbit slow and cardiac troponin I are more similar than either is to the rabbit fast troponin I and also that the C-terminal halves are correspondingly more different, but it is doubtful if this is significant.

The only other component of the troponin complex for which comparative sequence data are available are troponin C. The sequences of three fast skeletal muscle troponin C components have been published, from rabbit⁸, chicken²⁰ and human muscle²¹ and also the sequence of one cardiac troponin C from beef²². The difference between the rabbit and chicken proteins is 11% while there is only one difference between the rabbit and human proteins. All three fast muscle proteins differ from the cardiac protein by between 30% and 35%. This suggests that troponin C is a somewhat more conservative protein than troponin I.

In a comparison of the sequences the most striking feature is the extra 26 residues which the cardiac protein has at the N-terminal end which are not shared by the proteins from either fast or slow muscle. Apart from the N-terminal region the greatest differences are found in the region of residues 181–184 where one, two and three deletions have to be introduced into the rabbit slow, cardiac and fast muscle sequences respectively to fit them to the chicken fast muscle sequence. There is no discernable pattern in the sequence in this region and this is made more curious by the fact that the sequence on either side of this region is invariant. The sequence of the rabbit slow muscle protein includes a proline at residue 182 and this region may represent a bend in the polypeptide chain where mutations can be tolerated without affecting the overall three-dimensional structure.

Predictions of the secondary structure have been made^{23,24} and give values of about 45% α -helix and 10% β -sheet in rabbit skeletal troponin I whereas values determined from circular dichroic and optical rotary dispersion methods give values of 29% and 20% for α - and β -structure respectively²⁵, which are somewhat at variance with the predictions. There are two regions where concentrations of proline residues are found—in the regions of residues 76 and 137. The predicted structures emphasise that these are likely to be associated with bends in the structure, dividing the molecule into three roughly equal sections. These considerations, together with the highly polar nature of the proteins suggest that they exist as fairly open structures. In each case the charged amino acids account for some 40% of the residues and each protein possesses a net positive charge, this being 8 and 9 respectively in the case of rabbit and chicken fast muscle troponin I and 14 and 18 in the case of rabbit cardiac and slow muscle troponin I. The higher charge on the slow muscle protein is a consequence not only of an increased number of lysine and arginine residues but also of a decrease in the number of

aspartic and glutamic acid residues, there is not however any marked difference in the charge distribution in each molecule.

Studies on the phosphorylation of the rabbit fast and cardiac muscle troponin I show that a number of sites can be phosphorylated using phosphorylase kinase and a cyclic AMP-dependent protein kinase^{26–28}. The position of these sites is shown in Fig. 3.

The major phosphorylation sites in the rabbit fast muscle protein are at residues 37 and 146 (threonine 11 and serine 117 in the fast muscle troponin I sequence¹¹). Phosphorylation at both these positions is blocked in the presence of troponin C and it is doubtful if phosphorylation at these sites is of any physiological significance although they might play a part in the initial assembly of the troponin complex. In each of the proteins there is either a serine or a threonine present at positions 37 and 146 and the sequence around these residues shows considerable homology, it is possible that each of these positions may be phosphorylated in appropriate conditions, but with the exception of serine 146 (ref. 28) in cardiac troponin I this has not yet been demonstrated. The major phosphorylation site in cardiac troponin I is at serine 20^{28,30}, which is in the extra N-terminal part of the sequence. This residue is rapidly phosphorylated by cardiac cyclic AMP-dependent protein kinase and since it has been shown by Solaro *et al.*³⁰ to be phosphorylated when the perfused heart is stimulated with adrenaline, it is tempting to speculate that the part of the sequence which is unique to the cardiac protein is of functional importance in the regulation of cardiac muscle contraction.

Fig. 2 Difference matrices for the amino acid sequences of troponin I components. The numbers shown are % differences between sequences, ignoring extra terminal residues and counting deletions as an extra amino acid. C.F., chicken fast muscle; R.F., rabbit fast muscle; R.S., rabbit slow muscle; R.C., rabbit cardiac muscle. *a*, Comparison of whole proteins; *b*, comparison of sequences up to residue 119; *c*, comparison of sequences from residue 120 to the C-terminus.

	CF	RF	RS	RC
CF	0	18	39	42
PF	18	0	40	41
RS	39	40	0	40
RC	42	41	40	0
<i>a</i>				
	CF	RF	RS	RC
CF	0	27	55	56
RF	27	0	55	54
RS	55	55	0	48
RC	56	54	48	0
<i>b</i>				
	CF	RF	RS	RC
CF	0	10	22	28
RF	10	0	25	27
RS	22	25	0	33
RC	28	27	33	0
<i>c</i>				

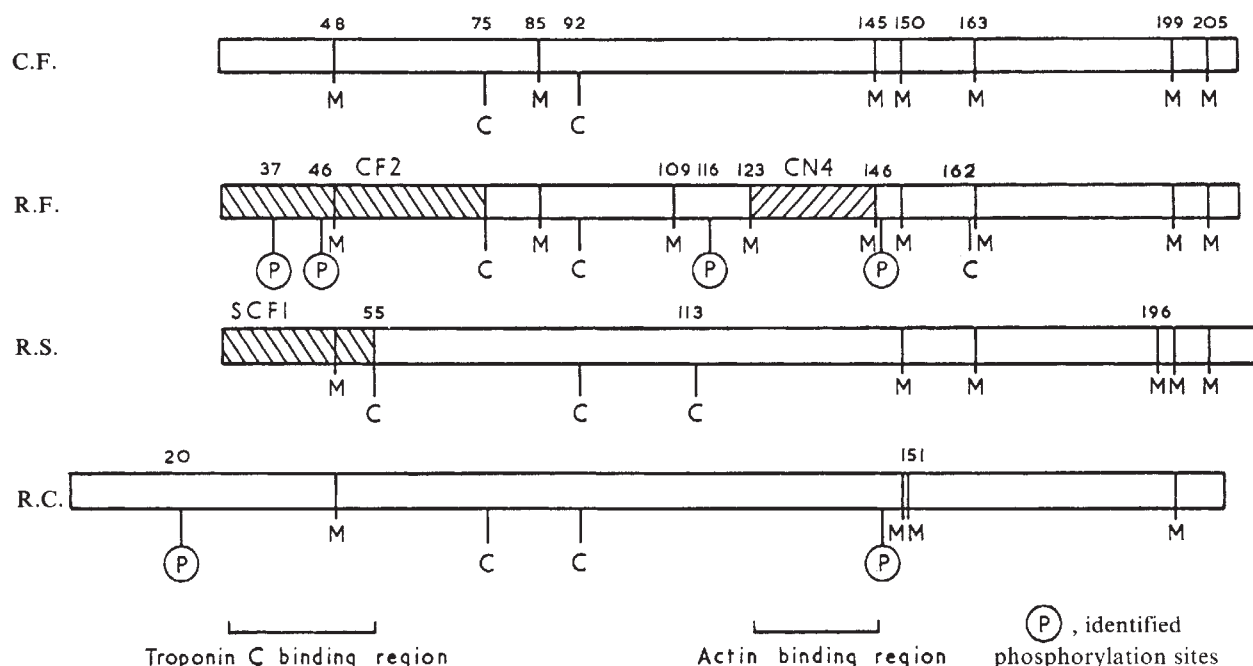


Fig. 3 Schematic representation of troponin I proteins from chicken fast (C.F.), rabbit fast (R.F.), rabbit slow (R.S.) and rabbit cardiac (R.C.) muscle. C, cysteine residues; M, methionine residues. The numbering system is that for the cardiac protein as shown in Fig. 1. In the R.F., R.S. and R.C. proteins only unique identified phosphorylation sites, C and M residues are numbered. The shaded areas indicate the fragments which interact with troponin C (CF2 and SCF1) and actin (CN4) as discussed in the text.

Biological activity of troponin I peptides

Two regions of the rabbit fast muscle troponin I molecule have been implicated in interactions with other components of the thin filament by Syska *et al.*³¹. The cyanogen bromide fragment CN4 (residues 124–145) has been shown to inhibit the actomyosin ATPase, this inhibition being enhanced by tropomyosin and reaching up to 70% of the inhibition obtained with whole troponin I and tropomyosin. This fragment has also been shown by affinity chromatography to bind to actin and also, to some extent, to troponin C. The troponin C binding region has been located at the N terminus of the molecule, partly by the inhibition of threonine-37 phosphorylation²⁹ and more particularly by the demonstration of the binding of the cysteine cleavage fragment CF2 (residues 27–74) to troponin C, both by affinity chromatography and gel electrophoresis³¹.

A consideration of the methionine positions in each of the proteins (Fig. 3), shows that methionine 123 is unique to rabbit fast muscle troponin I. It has probably arisen fairly late in the evolution of fast muscle troponin I and is therefore unlikely to occur in proteins from other sources. It is thus not possible to prepare analogous fragments from the other proteins in order to compare their inhibitory activity. Inspection of the sequence in this region (Fig. 1) shows, however, that it is one of the most strongly conserved in all of the four proteins. With the exception of residue 124, where all the proteins differ, the replacements in general involve strongly conservative hydrophobic residues. There is also a Thr/Pro replacement at position 138 in the cardiac protein and a possible extra leucine at position 141, which has been discussed above. It seems unlikely that any of these changes would affect the inhibitory activity significantly and if there are real differences in this their cause must be sought in other parts of the troponin I structure. This constancy of sequence at the actin binding site is perhaps not surprising in view of the very conservative nature of actin itself.

It is difficult to define the troponin C binding site as precisely as that for actin binding as appropriate fragments are not available. The cyanogen bromide fragment CN5

(residues 27–48) from the rabbit fast muscle protein shows some troponin C binding activity but is not nearly as effective as CF2. It is probable that the binding site spans methionine 48 but may not extend as far as cysteine 75. The rabbit slow muscle protein has an extra cysteine at residue 55 and the fragment produced by cleavage at this residue has been shown to bind to a troponin C–Sepharose column¹⁶, but owing to incomplete cleavage the amount of fragment obtained was small and no quantitative estimate of binding could be obtained. On the basis of homology the most likely binding site is between residues 36 and 52 but even in this region there is considerable variation.

As noted above, the inhibitory peptide (residues 124–145) has some binding affinity for troponin C. Weeks and Perry³² have shown that two of the large cyanogen bromide peptides from troponin C will neutralise the inhibition of the actomyosin ATPase by troponin I and will also inhibit phosphorylation at residues 37 and 146. It is not clear, however, if these observations indicate a second specific binding site for troponin C on the troponin I molecule, in the same position as the actin binding site, or whether they are due to strong ionic interactions between highly acidic peptides and a highly basic region in troponin I.

It is clear from the comparison of the sequences that the genes coding for the three types of striated muscle troponin I diverged at much the same time in evolutionary terms and well before the divergence of birds and mammals.

The overall rate of evolution has been significantly greater than that for troponin C and this must be a reflection of the differing constraints placed on the tertiary structures. Troponin C has four Ca^{2+} binding sites and also interacts with two other components of the thin filament. When one of the Ca^{2+} binding sites is lost, as in the cardiac protein²², the sequence in this region is found to vary considerably. Troponin I has only two protein binding sites and thus a higher rate of mutation is to be expected. But, as shown above, the N- and C-terminal parts of the molecule have changed at rather different rates—the C-terminal half having much the same rate of evolution as troponin C whereas the N-terminal half has changed at nearly twice this rate.

The constancy of the C-terminal parts is explicable in terms of binding to the strongly conserved actin molecule and this would suggest that the mechanism for the inhibition of the actomyosin ATPase is the same for each of the three types of striated muscle troponin I. The much greater variability in the N-terminal part of troponin I and in particular in the troponin C binding region is somewhat surprising in view of the conservative nature of troponin C itself and this may reflect subtleties in the interaction of troponin I with troponin C which may only be appreciated from an understanding of the three-dimensional structure of the troponin complex.

We thank Professor S. V. Perry and Dr A. J. G. Moir for discussions and the MRC for financial support.

Received 2 August; accepted 20 October 1977.

1. Ebashi, S., Endo, M. & Ohtsuki, I. *Q. Rev. Biophys.* **2**, 351–384 (1969).
2. Weber, A. & Murray, J. M. *Physiol. Rev.* **653**, 612–673 (1973).
3. Wilkinson, J. M., Perry, S. V., Cole, H. A. & Trayer, I. P. *Biochem. J.* **127**, 215–228 (1972).
4. Ebashi, S., Ohtsuki, I. & Mihashi, K. *Cold Spring Harb. Symp. quant. Biol.* **37**, 215–223 (1972).
5. Greaser, M. L. & Gergely, J. *J. biol. Chem.* **248**, 2125–2133 (1973).
6. Potter, J. & Gergely, J. *Biochemistry* **13**, 2697–2703 (1974).

7. Head, J. F. & Perry, S. V. *Biochem. J.* **137**, 145–154 (1974).
8. Collins, J. H., Potter, J. D., Horn, M. J., Wilshire, G. & Jackman, N. *FEBS Lett.* **36**, 268–272 (1973).
9. Elzinga, M., Collins, J. H., Kuehl, W. M. & Adelstein, R. S. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2687–2691 (1973).
10. Stone, D., Sodek, J., Johnson, P. & Smillie, L. B. *Proc. FEBS Meet.* 9th 31, 125–136 (1975).
11. Wilkinson, J. M. & Grand, R. J. A. *Biochem. J.* **149**, 493–496 (1975).
12. Pearlstone, J. R., Carpenter, M. R., Johnson, P. & Smillie, L. B. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1902–1906 (1976).
13. Greaser, M. L., Yamaguchi, M., Brekke, C., Potter, J. & Gergely, J. *Cold Spring Harb. Symp. quant. Biol.* **37**, 235–244 (1972).
14. Tsukui, R. & Ebashi, S. *J. Biochem.* **73**, 1119–1121 (1973).
15. Syska, H., Perry, S. V. & Trayer, I. P. *FEBS Lett.* **40**, 253–257 (1974).
16. Grand, R. J. A. & Wilkinson, J. M. *Biochem. J.* **167**, 183–192 (1977).
17. Grand, R. J. A., Wilkinson, J. M. & Mole, L. E. *Biochem. J.* **159**, 633–641 (1976).
18. Frank, G. & Weeds, A. G. *Eur. J. Biochem.* **44**, 317–334 (1974).
19. Collins, J. H. *Nature* **259**, 699–700 (1976).
20. Wilkinson, J. M. *FEBS Lett.* **70**, 254–256 (1976).
21. Romero-Herrera, A. E., Castillo, O. & Lehmann, H. J. *molec. Evol.* **8**, 251–270 (1976).
22. van Eerd, J.-P. & Takahashi, K. *Biochem. biophys. Res. Commun.* **64**, 122–127 (1975).
23. Wilkinson, J. M. & Grand, R. J. A. *Proc. FEBS Meet.* 9th 31, 137–144 (1975).
24. Ptiitsyn, O. B., Finkelstein, A. V. & Lim, V. I. *Proc. FEBS Meet.* 9th 31, 145–160 (1975).
25. Wu, C.-S. C. & Yang, J. T. *Biochemistry* **15**, 3007–3014 (1976).
26. Huang, T. S., Bylund, D. B., Stull, J. T. & Krebs, E. G. *FEBS Lett.* **42**, 249–252 (1974).
27. Moir, A. J. G., Wilkinson, J. M. & Perry, S. V. *FEBS Lett.* **42**, 253–256 (1974).
28. Moir, A. J. G. & Perry, S. V. *Biochem. J.* **167**, 333–343 (1977).
29. Perry, S. V. & Cole, H. A. *Biochem. J.* **141**, 733–743 (1974).
30. Solaro, R. J., Moir, A. J. G. & Perry, S. V. *Nature* **262**, 615–617 (1976).
31. Syska, H., Wilkinson, J. M., Grand, R. J. A. & Perry, S. V. *Biochem. J.* **153**, 375–387 (1976).
32. Weeks, R. A. & Perry, S. V. *Biochem. Soc. Trans.* **5**, 1391–1392 (1977).

letters to nature

Discovery of an X-ray QSO

WE report here the discovery of an X-ray emitting QSO, the first to be initially identified from X-ray observations. Previously, the only QSO known to be an X-ray source was 3C273 (refs 1 and 2). The new QSO has been found within a 40'' error circle established by the SAS-3 X-ray Observatory. The SAS-3 error circle lies within the Ariel V error box (~ 0.2 square degrees) for the source 2A 2251-179 (ref. 3). Following the convention for optical QSOs, we have designated the optical object MR2251-178. The X-ray luminosity (2–11 keV) of this object, presently $\sim 5 \times 10^{44}$ erg s $^{-1}$, has been as large as $\sim 1.6 \times 10^{45}$ erg s $^{-1}$ (in 1975) and exceeds the optical luminosity by a factor of ~ 10 . Among known compact X-ray sources, only 3C273 has a greater luminosity. Also, in radio observations at 4,885 MHz with the NRAO Very Large Array (VLA), we have discovered a point-like radio source coincident with MR2251-178.

A $12^\circ \times 12^\circ$ field near 2A2251-179 was observed with the rotating modulation collimator (RMC) system (ref. 4) on the SAS-3 X-ray Observatory from 11.1 August 1977 to 13.5 August 1977. Data from 36 orbits (effective integration time of 9.7×10^4 s) have been analysed.

Detection of a source within the error box for 2A2251-179 was achieved in each of the two independent RMC detectors. The statistical significance of the detection was 6.2σ in the 213 FWHM detector (2–11 keV) and 5.7σ in the 415 FWHM detector (2–11 keV). In Table 1 the results of these two independent detections are given, including spectral data for the

two energy channels in each of the two detectors. While the spectral shape cannot be uniquely determined, the data are consistent with a power law having a photon number index of 1.5 ± 0.5 (assuming $E_0 \approx 2$ keV; ref. 5). The flux is $2.5 \pm 0.4 \times 10^{-11}$ erg cm s $^{-1}$ (2–11 keV). If we combine the two independent position determinations according to the technique described by Doxsey *et al.*⁴ we obtain the following position for the X-ray source (in subsequent catalogues, the X-ray source will be designated 2S2251-178).

$$\alpha(1950) = 22 \text{ h } 51 \text{ min } 25.4 \text{ s}, \quad \delta(1950) = -17^\circ 50' 40'' \\ \text{error circle radius (90\% confidence): } 40''$$

In Fig. 1, we show the Ariel V error box for 2A2251-179 (0.187 square degrees in area; ref. 3) and the refined SAS-3 error circle reported here. Both are superimposed on a Palomar Observatory Sky Survey (POSS) blue plate. Figure 1 inset shows only the SAS-3 error circle, but with a scale factor twenty times larger. Within the SAS-3 error circle, there are only 2 conspicuous objects ($m_v \lesssim 20$) visible on the POSS plates. They are denoted '1' and '2' on the inset. Object 1 possibly is a compact elliptical galaxy, whereas Object 2 is starlike. There is no marked difference between the red and blue POSS images of either object. Coordinates for these two objects were measured on POSS glass plates at Kitt Peak National Observatory (KPNO) by S. Kleinmann. Plates from the Tololo-Michigan Objective Prism Survey (previously taken with the 61 cm Curtis Schmidt telescope at Cerro Tololo; refs 6 and 7) were then examined. On these plates, the spectrum of Object 1 is obscured in part by that of Object 2. In the most remarkable spectrum of Object 2, bright emission lines from the following species are evident: H β , [OIII] $\lambda 3727$, [NeIII] 3869, H γ , and probably [NeV] $\lambda 3426$ and HeII $\lambda 4686$. Using 4 lines in the 3,500–5,350-Å interval, we determined a redshift of $z = 0.0680 \pm 0.003$. This implies a distance of 410 Mpc, assuming $H_0 = 50$ km s $^{-1}$ Mpc $^{-1}$. Based on the diameter of its POSS image, the apparent magnitude of Object 2 is $B \sim +16$.

Table 1 X-ray observations of MR2251-178 (2A2251-179) by SAS-3

Collimator f.w.h.m.	Energy band (keV)	Counting rate (s $^{-1}$)
213	2–6	0.45 ± 0.08
	6–11	0.14 ± 0.05
415	2–6	0.30 ± 0.06
	6–11	0.13 ± 0.05

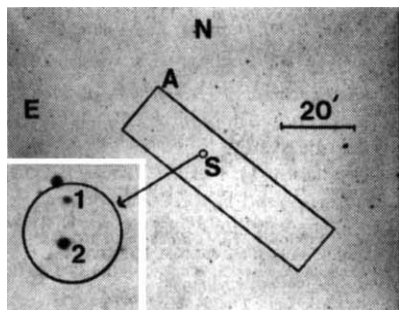


Fig. 1 The QSO MR2251-178 and the X-ray source 2A2251-179. 'A' and 'S' denote the Ariel V error box and the SAS-3 error circle, respectively. The inset shows the SAS-3 error circle enlarged by a factor of 20. Object 2 is MR2251-178. Photographs are taken from the Palomar Sky Survey Plates National Geographic Society).

The B luminosity of the object is then $\sim 5 \times 10^{43}$ erg s^{-1} , while the X-ray luminosity is $\sim 5 \times 10^{44}$ erg s^{-1} (2-11 keV). The position of the optical object (designated MR2251-178) is

$$\alpha(1950) = 22 \text{ h } 51 \text{ min } 25.89 \text{ s} \pm 0.03 \text{ s}, \\ \delta(1950) = -17^\circ 50' 54''.2 \pm 0''.5$$

The coordinates of the centroid of Object 1 were also measured at KPNO. They are

$$\alpha(1950) = 22 \text{ h } 51 \text{ min } 25.82 \text{ s} \pm 0.03 \text{ s}, \\ \delta(1950) = -17^\circ 50' 15.3'' \pm 0.5''$$

A search for radio emission from MR2251-178 was carried out at a frequency of 4,885 MHz using 6 antennas of the NRAO Very Large Array (VLA) in New Mexico. Operating with 50 MHz bandwidth and system temperatures of roughly 60 K, the six 25-m antennas form an aperture synthesis system capable of easily detecting 1 mJy sources in a $5'$ field in 1 h. With the six antennas arranged in a linear array, 7.7 km in length, the resolution of the system is <1 arc s. Aperture synthesis maps of the field containing MR2251-178 made with 1.8 and 2.4 h of data on 4 and 6 October, 1977, independently show the presence of a dominant 3.2 ± 0.3 mJy point source at a position of

$$\alpha(1950) = 22 \text{ h } 51 \text{ min } 25.90 \text{ s} \pm 0.04 \text{ s}, \\ \delta(1950) = -17^\circ 50' 53.5'' \pm 1''$$

No variability was noted between the two dates and the source was <1 arc s in extent.

In the field containing MR2251-178, there is also an extended radio source of very low surface brightness at

$$\alpha(1950) = 22 \text{ h } 51 \text{ min } 25.8 \text{ s} \pm 0.8 \text{ s}, \\ \delta(1950) = -17^\circ 50' 20'' \pm 10''$$

This source has a size of the order of $30''$ and contains at least 10 mJy of radio flux at 4,885 MHz. Because of the elongation of the synthesised beam, it is not possible at the present time to discuss the detailed structure of this extended source; however, considering its position and size, it is very likely associated with Object 1 in Fig. 1.

Based on the positional coincidence (~ 40 arc s), we regard the identification of the X-ray source 2A2251-179 with the optical object MR2251-178 as virtually certain. The association of the VLA radio point source with MR2251-178 is certain (~ 1 arc s positional coincidence). The classification of MR2251-178 as a QSO is based upon its stellar appearance on POSS plates, its redshift, and its high luminosity. It is possible that it could be an extreme Seyfert galaxy. This, however, would require that its luminosity be greater than

that of any known X-ray-emitting Seyfert galaxy (refs 8 and 9). 2A2251-179 is reported to vary by a factor of ~ 10 in the 2A catalogue (ref. 3). At the maximum brightness reported, the 2-11 keV luminosity would be $\sim 1.6 \times 10^{45}$ erg s^{-1} . This exceeds the maximum value reported for the Seyfert galaxy of highest reported X-ray luminosity, 3C390.3 (refs 10 and 11).

Among the known compact X-ray sources, only 3C273 has a maximum luminosity greater than MR2251-178 (ref. 10). The ratio of X-ray to optical emission we have observed for MR2251-178 is $\sim 10:1$ compared to a ratio of $\sim 1:1$ in 3C273 and typical X-ray Seyferts (ref. 10). Thus, judging from available data, the energy budget of MR2251-178 is dominated by its X-ray emission. Accurate assessments of its bolometric luminosity must await infrared and high energy (>10 keV) X-ray measurements.

The redshift of MR2251-178 would place it among the closer QSOs. In all, there are only 5 (out of 633) QSOs in the list of Burbidge, Crown, and Smith (ref. 12) which have smaller redshifts. At the redshift distance of MR2251-178, the projected separation between it and Object 1 is ~ 75 kpc.

The presence of the extended radio source near Object 1 may suggest a possible link to MR2251-178. A redshift measurement of Object 1 would clarify this possibility.

The optical and radio properties of MR2251-178 are such that it is unlikely that it would have been detected in pre-Tololo surveys. In the optical, it would likely have been passed over as an uninteresting field star, while in the radio, it is too faint to have been detected in any standard survey to date. Thus, there could be many such sources. Even if their number density were as great as that of Seyfert galaxies ($\sim 0.5\%$ of all galaxies; ref. 13), they could have been overlooked in optical searches, since the objects would be ~ 2.5 mag fainter for a given X-ray luminosity than are the Type I Seyfert galaxies studied by Ariel V (ref. 8), of which the optically faintest was $m_v \sim 14.5$. In view of the hard spectrum of MR2251-178 in the X-ray region (comparable in index to the diffuse X-ray background; ref. 14), the general population of such objects could make a substantial contribution to the X-ray background.

We acknowledge support from the staffs of the Center for Space Research at MIT and the Goddard Space Flight Center of NASA. We thank H. Bradt, C. Canizares, S. Kleinmann, R. Kelley, W. Lewin, J. McClintock, and B. Rossi of MIT, and B. Cooke, M. Elvis and J. Pye of the Ariel V Group, University of Leicester for helpful discussions and assistance. The National Radio Astronomy Observatory is operated by Associated Universities, Inc. under contract with the US NSF. This work was supported in part by NASA under contract NAS5-11450.

G. R. RICKER
G. W. CLARKE
R. E. DOXSEY
R. G. DOWER
J. G. JERNIGAN

Department of Physics and Center for Space Research,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

J. P. DELVAILLE

Smithsonian Astrophysical Observatory,
Cambridge, Massachusetts 02138

G. M. MACALPINE

Department of Astronomy,
University of Michigan,
Ann Arbor, Michigan 48109

R. M. HJELLMING

National Radio Astronomy Observatory,
Socorro, New Mexico 87801

Received 12 October; accepted 24 October 1977.

1. Bowyer, S., Lampton, M., Mack, J. & de Mendonca, F. *Astrophys. J. Lett.* **161**, L1-L8 (1970).
2. Kellogg, E. et al. *Astrophys. J. Lett.* **165**, L49-L54 (1971).
3. Cooke, B. A. et al. *Mon. Not. R. astr. Soc.* (in the press).
4. Doxsey, R. F. et al. *Nature* **269**, 112-116 (1977).
5. Schnopper, H. W. et al. *Astrophys. J. Lett.* **215**, L7-L11 (1977).

6. MacAlpine, G. M., Smith, S. B. & Lewis, D. W. *Astrophys. J. (Suppl.)* **34**, 95–100 (1977).
7. Smith, M. G. *Astrophys. J.* **202**, 591–595 (1975).
8. Elvis, M. et al. *Mon. Not. R. astr. Soc.* (in the press).
9. Tananbaum, H. et al. *Astrophys. J. Lett.* (in the press).
10. Culhane, J. L. *Q. J. R. Astr. Soc.* (in the press).
11. Charles, P. A., Longair, M. S. & Sanford, P. W. *Mon. Not. R. astr. Soc.* **170**, 17p–22p (1975).
12. Burbidge, G. R., Crown, A. H. & Smith, H. E. *Astrophys. J. Suppl.* **33**, 113–188 (1977).
13. Huchra, J. & Sargent, W. L. W. *Astrophys. J.* **186**, 433–444 (1973).
14. Schwartz, D. & Gursky, H. in *X-Ray Astronomy* (eds R. Giacconi & H. Gursky) 359–388 (Reidel, Dordrecht, 1974).

Observation of X-ray eclipses from LMC X-4

THERE are six known X-ray sources in the Magellanic Clouds: five in the Large Cloud and one in the Small Cloud^{1,2}. Until recently, only SMC X-1 was known to be in a binary system³. Chevalier and Ilovaisky⁴ have now shown that the optical emission of the suggested counterpart of LMC X-4 (ref. 5) exhibits ellipsoidal light variations with an inferred binary period of 1.408 d. Previous X-ray observations have indicated that LMC X-4 is highly variable on time scales from minutes to months^{2,6–8}. We report here the detection of several eclipses of LMC X-4, which confirms it as the second known extragalactic X-ray binary. We also discuss briefly the inferred value of the mass of the compact X-ray star.

LMC X-4 was observed on four occasions with the X-ray detectors on SAS-3. The first observation⁶ was made with the Rotation Modulation Collimator system (RMC) during 20.5–26.4 February 1976 (UT). During the interval 15.8–17.6 October 1976 (UT), LMC X-4 was scanned twice per satellite orbit

with the Right Slat detector system^{9,10}. The Y-axis detectors^{9,10} were then pointed continuously at LMC X-4 during the intervals 22.2–23.2 and 23.99–24.32 May 1977 (UT). During the first of these pointed observations, SAS-3 was operated with a special telemetry mode whereby the data from the 6–12 keV channel of the Horizontal Tube system^{9,10} were recorded with a temporal resolution of 8 ms (ref. 11). For the second pointed observation, as well as the scanning and RMC observations, the normal 0.83 s time resolution of the satellite was used.

LMC X-4 was extremely weak compared to the non-X-ray background during most of our observations. We were, therefore, unable to obtain accurate intensities for the source. Instead, we assigned a qualitative description to the observed intensities as

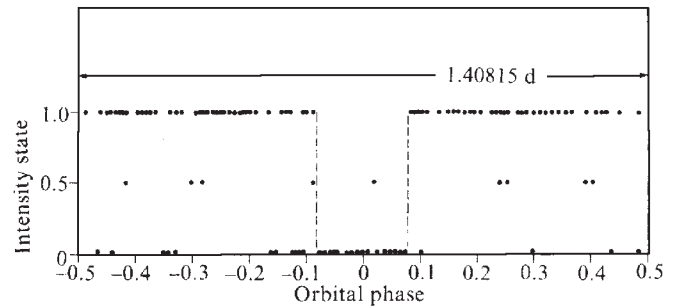


Fig. 1 Orbital X-ray light curve for LMC X-4. The 'digitised' X-ray intensities for LMC X-4, given in Table 1, are folded modulo the trial orbital period 1.40815 d. The apparent eclipse of the X-ray source has a duration of 0.23 ± 0.02 d and is centred at JD 2443068.01 ± 0.02 .

Table 1 Observed intensities for LMC X-4

Time*	Intensity† state	Time	Intensity state	Time	Intensity state
28.952	1	33.747	0	268.708	0
29.410	1	33.811	1/2	268.823	1
29.477	1	33.878	1	268.839	1
29.607	1	33.942	1	268.888	1
29.871	1	34.009	0	268.954	1
29.937	0	34.073	0	269.020	1
30.132	1	34.140	0	269.035	1
30.198	1	34.207	0	269.085	1
30.265	1	34.271	0	269.151	1
30.331	1	34.338	0	269.167	1
30.395	1	267.375	1	269.216	0
30.463	1	267.398	1	269.232	1
30.527	1	267.440	1/2	269.282	0
30.593	1	267.538	0	269.560	1
30.791	1	267.561	0	269.576	1
31.053	1	267.635	1	269.695	1
31.120	1	267.658	1	269.761	1
31.184	1	267.725	1	269.826	1
31.251	0	267.835	1	269.871	1
31.315	0	267.857	0	269.937	1
31.382	0	267.900	1/2	270.002	1/2
31.514	0	267.923	0	270.021	1
31.654	1	268.034	0	270.068	1
31.711	1	268.051	1/2	485.773	1
31.775	1	268.100	0	485.839	1
31.842	1	268.117	0	485.970	1
31.906	1	268.165	1	486.101	1
31.973	1/2	268.182	1	486.167	1
32.105	1	268.231	1	486.298	0
32.169	0	268.248	1	486.364	0
32.236	1	268.297	1	486.429	0
32.301	1	268.314	1	486.495	1
32.367	1	268.362	1/2	486.626	1
32.433	1/2	268.379	1/2	486.757	1
32.497	1	268.428	1	487.552	1
32.760	0	268.445	0	487.618	0
32.827	0	268.493	1	487.683	0
32.891	0	268.511	1	487.749	0
33.549	1	268.576	1	487.814	1
33.616	0	268.625	1		
33.680	1	268.642	0		

*JD—2442800.0

†See text for definition.

follows: 1 when the intensity was at least 3σ above the background, 0 when it was less than 1σ above the background and 1/2 when we could not determine definitively whether the source was 'on' or not. We then compiled a list of these intensity states, one for each scan during the scanning observation and one for every satellite orbit during the other observations. The results are presented in Table 1.

It is evident from Table 1 that LMC X-4 was 'off' for ≥ 5 h on six separate occasions. This behaviour is similar to the eclipses observed in other eclipsing X-ray binaries¹². Furthermore, the last two 'off' states occurred very close to the times when the unseen secondary was predicted to be behind the primary⁴. Unfortunately, the orbital period determined from the optical data is not sufficiently accurate to determine whether the earlier X-ray 'off' states occurred at a predicted eclipse time also.

The X-ray 'intensities' from all four observations were folded modulo trial binary orbital periods between 1.404 d and 1.412 d in steps of 0.0001 d. The period range is chosen to be within two standard deviations of the optical period⁴. Figure 1 displays the 'light' curve for the trial orbital period of 1.40815 d. An eclipse is clearly discernible with a width of 0.23 ± 0.02 d. (It is difficult to assign a formal confidence to the errors quoted in this paragraph; however, we regard these errors as conservative limits.) When the data were folded at periods different from 1.40815 d by more than 0.00020 d, the resulting 'eclipses' were significantly shorter and eventually disappeared. We conclude that we have observed an X-ray eclipsing phenomenon with the above binary period, and with an eclipse centre at JD 2443068.01 ± 0.02 . This confirms the proposed optical identification of LMC X-4.

A search for X-ray pulsations in the data taken with 8 ms and 0.83 s time resolution was carried out. No significant peaks in the Fourier power density spectrum were observed. From these studies we place an upper limit to pulsing of $\sim 25\%$ of the steady flux with pulse periods ranging from ~ 16 ms to 400 s. We note that if LMC X-4 had a pulsation fraction and pulse shape similar to that of SMC X-1 (refs 13, 14) it would have been only marginally detectable as a pulsar.

From the velocity amplitude of the He(II) emission lines in the LMC X-4 binary system ($475 \pm 25 \text{ km s}^{-1}$; ref. 4), we estimate a lower limit (2σ) to the mass function:

$$f(M) = \frac{m_{\text{opt}}^3 \sin^3 i}{(m_{\text{opt}} + m_x)^2} = \frac{4\pi^2 (a_x \sin i)^3}{G P_{\text{orb}}^2} \gtrsim 11.2 M_{\odot}$$

where m_{opt} and m_x are the masses of the companion star and X-ray star, respectively, i is the orbital inclination angle, $a_x \sin i$ the projected semi-major axis of the orbit of the X-ray star, and P_{orb} the orbital period. To establish the limit on the mass function we have assumed that the He(II) emission arises from a region between the two stars⁴. We further define the mass ratio of the two stars as $q = m_x/m_{\text{opt}}$.

Figure 2 is a plot of contours of constant m_x in the q/i plane for a mass function $f(M) = 11.2 M_{\odot}$. The upper limit⁴ on the velocity of the optical companion [$\sim 79 \text{ km s}^{-1}$ (2σ)] implies that the mass ratio, q , is smaller than ~ 0.17 . Chevalier and Ilovaisky⁴ argued that i must be greater than $\sim 70^\circ$, because of the large amplitude of the ellipsoidal light variations. The measured X-ray eclipse half-angle, θ_e , is between 27° and 31° . If we assume that the optical companion fills at least 90% of its Roche lobe, then the above limits on q, i , and θ_e restrict m_x to be greater than $\sim 1.1 M_{\odot}$. (For a more complete discussion of this method of mass estimation, see refs 15, 16.) As the value of $f(M)$ we adopted is only a lower limit, we cannot provide a meaningful upper limit to m_x . The shaded region of Fig. 2 shows the allowable range of values of m_x for the various constraints imposed. If $f(M)$ is larger than $11.2 M_{\odot}$, the contours of constant m_x will shift to the lower right of the q/i plane, thereby increasing the lower limit on m_x . We note that if the data point with 'intensity' = 1/2, which occurs during

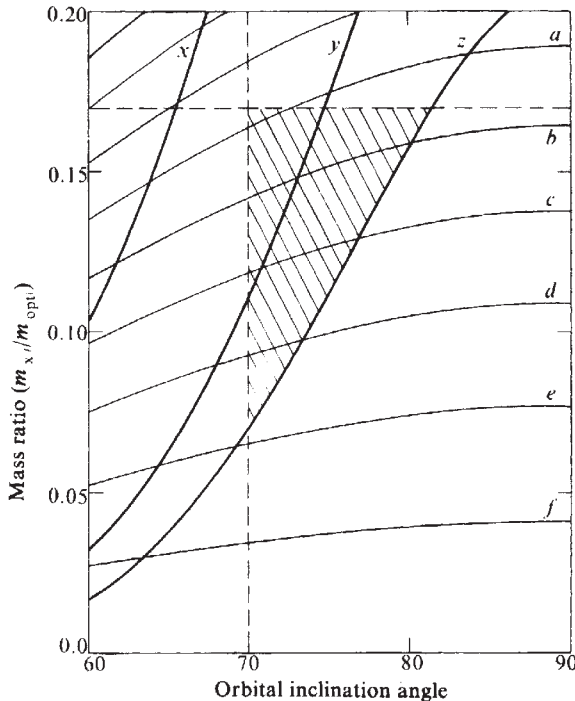


Fig. 2 Constraints on the mass, m_x , of the X-ray star in LMC X-4. The light solid curves are contours of constant m_x in the mass ratio/orbital inclination angle (q/i) plane (see text), for a mass function $f(M) = 11.2 M_{\odot}$: a-f are for $m_x = 3, 2.5, 2, 1.5, 1$ and $0.5 M_{\odot}$, respectively. The heavy solid curves are constraints imposed by the assumption of various orbital geometries and eclipse angles, θ_e (see ref. 14): x, y and z represent $(1.0 R_T, \theta_e 27^\circ)$, $(1.0 R_L, \theta_e 29^\circ)$, and $(0.9 R_L, \theta_e 31^\circ)$, respectively, $1.0 R_T$, $1.0 R_L$, and $0.9 R_L$ denotes cases where the optical companion star fills 100% or 90% of its critical tidal or Roche lobe, respectively. The other constraints on q and i are discussed in the text. The hatched region denotes the acceptable range of values of m_x for $f(M) = 11.2 M_{\odot}$. For larger values of $f(M)$ the contours of m_x shift down and to the right in the plot; all other features remain fixed.

the eclipse (Fig. 1), actually represents an 'on' state, the eclipse half-angle would then be reduced to $\sim 20^\circ$. This also has the effect of increasing the lower limit on m_x .

Since the completion of this work, we have become aware of a similar study¹⁷ by the X-ray group at the Mullard Space Science Laboratory. The values reported for the X-ray eclipse ($0.206 \pm 0.008 \text{ d}$) and the orbital period ($1.413 \pm 0.007 \text{ d}$) are in good agreement with the corresponding values quoted in the present work, though of differing precisions.

The authors wish to thank the SAS-3 group for their assistance during this study and, in particular, G. Jernigan and P. Joss for helpful discussions and P. Hertz for preparing the figures. S.R. acknowledges support from the Alfred P. Sloan Foundation. This work was supported in part by NASA.

F. LI

S. RAPPAPORT

Department of Physics and Center for Space Research,
Massachusetts Institute of Technology,
Cambridge, Massachusetts 02139

A. EPSTEIN

Center for Astrophysics,
Smithsonian Astrophysical Observatory,
Cambridge, Massachusetts 02138

Received 13 September; accepted 3 November 1977.

1. Giacconi, R. *et al. Astrophys. J. Suppl.* **27**, 37-64 (1973).
2. Markert, T. & Clark, G. W. *Astrophys. J. Lett.* **196**, L55-L58 (1975).
3. Schreier, E., Giacconi, R., Gursky, H., Kellogg, E. & Tananbaum, H. *Astrophys. J. Lett.* **178**, L71-L75 (1972).
4. Chevalier, C. & Ilovaisky, S. A. *Astr. Astrophys.* **59**, L9-L12 (1977).
5. Sanduleak, N. & Philip, A. G. D. *IAU Circ. No.* 3023 (1976).
6. Epstein, A. *et al. Astrophys. J.* **216**, 103-107 (1977).
7. Rapley, C. G. & Tuohy, I. R. *Astrophys. J. Lett.* **191**, L113-L116 (1974).
8. Griffiths, R. E. & Seward, F. D. *Mon. not. R. astr. Soc.* **180**, 75P-79P (1977).
9. Lewin, W. H. G. *et al. Mon. not. R. astr. Soc.* **177**, 93P-102P (1976).
10. Buff, J. *et al. Astrophys. J.* **212**, 768-773 (1977).
11. Primini, F. *et al. Astrophys. J. Lett.* **210**, L71-L74 (1976).
12. Blumenthal, G. R. & Tucker, W. H. A. *Rev. Astr. Astrophys.* **12**, 23 (1974) & refs therein.
13. Lucke, R., Ventis, D., Friedman, H., Fritz, G. & Shulman, S. *Astrophys. J. Lett.* **206**, L25-L28 (1976).
14. Primini, F., Rappaport, S. & Joss, P. C. *Astrophys. J.* (in the press).
15. Avni, Y. & Bahcall, J. N. *Astrophys. J.* **197**, 675-688 (1975).
16. Rappaport, S. & Joss, P. C. *Ann. N.Y. Acad. Sci.* (in the press).
17. White, N. E. & Davidson, P. J. *IAU Circ. No.* 3095 (1977).

Discovery of eclipsing nature of LMC X-4

THE Sanduleak-Philip optical candidate for LMC X-4 (ref. 1) has been reported by Chevalier and Ilovaisky² to be part of a 1.4-d binary system. Six days of data taken from an extended Ariel V observation of the LMC are shown in Fig. 1 and X-ray eclipses 5 h long can be clearly discerned every 1.4 d. The predicted optical phase zero times are, within the uncertainties, in good agreement with the observed mid-eclipse times, thus confirming the identification and making this the first X-ray eclipsing binary system to be found in the LMC.

The 3.75° FWHM field of view of the proportional counter spectrometer (PCS) was centred on LMC X-4 but also contained LMC X-1, LMC X-3 (ref. 3) and AO538-66 (refs 4, 5). The three latter sources plus contributions from the particle and X-ray backgrounds are responsible for the residual flux seen during LMC X-4 eclipses. The gaps in the data are caused by the satellite turning off during its night time, regions of high charged particle background and lost data. In spite of this, both eclipse ingress and egress were seen on two occasions and from these an orbital period of $1.413 \pm 0.007 \text{ d}$ was obtained (phase zero occurring on JD 2443342.614), in good agreement with the optical value of $1.408 \pm 0.002 \text{ d}$. Combining the X-ray and optical epochs (obtained 4 months apart) gives an improved period of $1.40838 \pm 0.00027 \text{ d}$.

Figure 2 shows one of the eclipses plotted with a time resolution of 64 s. Entry and exit occur on a timescale of minutes with an eclipse duration on the two occasions where it could be measured of $0.208 \pm 0.008 \text{ d}$ ($\theta_e = 26.2 \pm 1.1^\circ$). This may be combined with the available optical data to derive a mass for the secondary, but this is dependent both on the inclination of the system and the

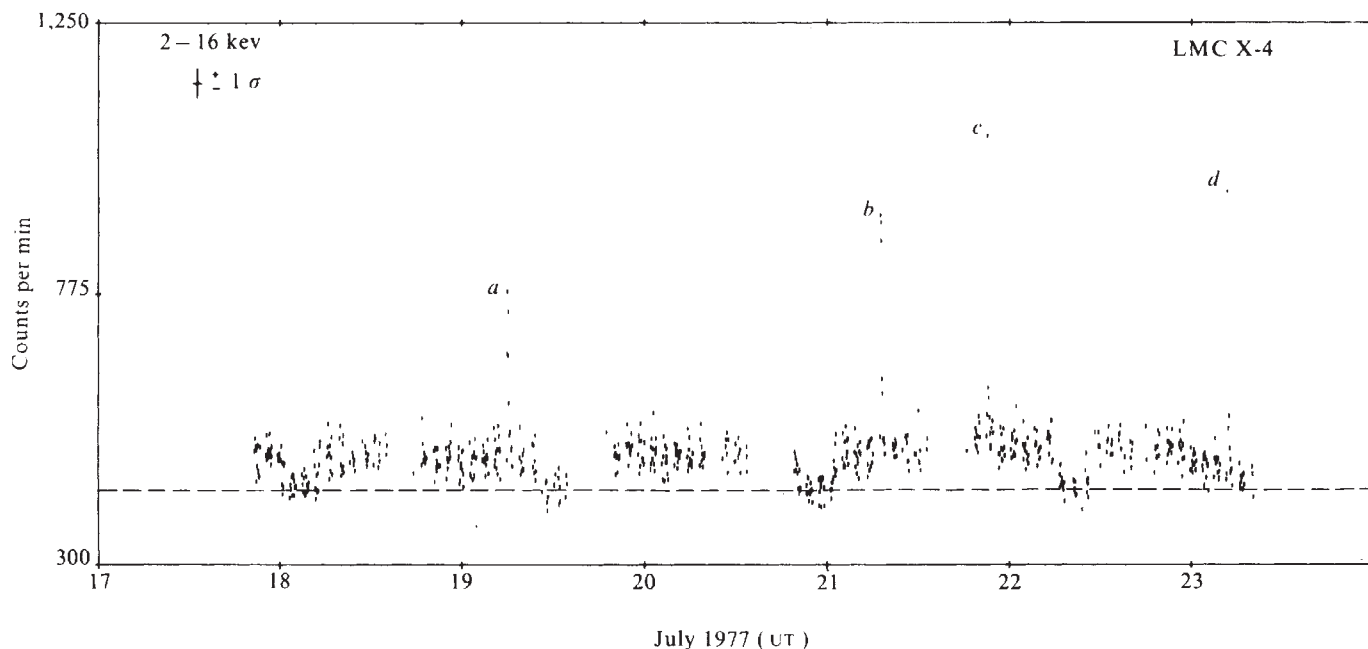


Fig. 1 The flux seen by the PCS integrated in 8 min time bins. The decline in flux during the last few hours can be attributed to the pointing position moving away from LMC X-4, LMC X-1 and LMC X-3.

assumed radius of the primary star. At present, these two parameters are not sufficiently constrained to improve significantly the estimate of $1-3 M_{\odot}$ already derived by Chevalier and Ilovaisky².

The X-ray data were unsuccessfully searched for evidence of periodicities between ~ 1 s and 40 min with an upper limit of

Four flares each lasting between 10 and 20 min were seen and they are shown in Fig. 3 plotted with the shortest time resolution available of 64 s. These 'outbursts' could have originated from any source in the field of view of the PCS. The PCS, however, is coaligned with a rotation modulation collimator which can obtain a source intensity averaged over one satellite orbit (typically ~ 50 min out of every 100) and combining two of the orbits when the flares occurred indicated a factor two increase in the intensity from LMC X-4 (G. Carpenter, personal communication). This suggests that the outbursts did originate from LMC X-4 and that the peak represents a source luminosity of 9×10^{38} erg s⁻¹. We note that Epstein *et al.*⁶ have reported 20-s

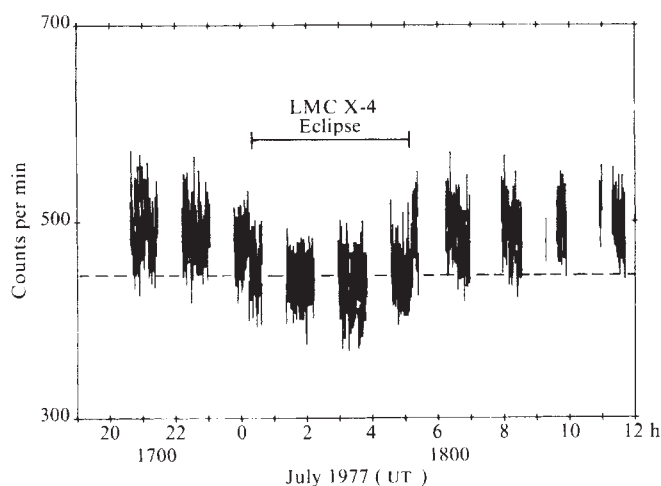


Fig. 2 One eclipse from LMC X-4 plotted with the finest time resolution available (64 s).

$\sim 20\%$ the mean flux. The overall peak luminosity of LMC X-4 was $\sim 3.4 \times 10^{-10}$ erg cm⁻² s⁻¹ in the 1.7 to 17.0 keV band which represents a total source output of 1×10^{38} erg s⁻¹. X-ray eclipses were only seen between 15 and 23 July, although the observations lasted from 29 June to 23 July and from 3 to 7 August. Six days of data obtained from 30 June onwards were folded modulo the orbital period and a two sigma upper limit to any residual flux above background (as defined during the predicted eclipse) of 1.8×10^{-11} erg cm⁻² s⁻¹ was found. This represents an upper limit to the source luminosity of 5×10^{36} erg s⁻¹.

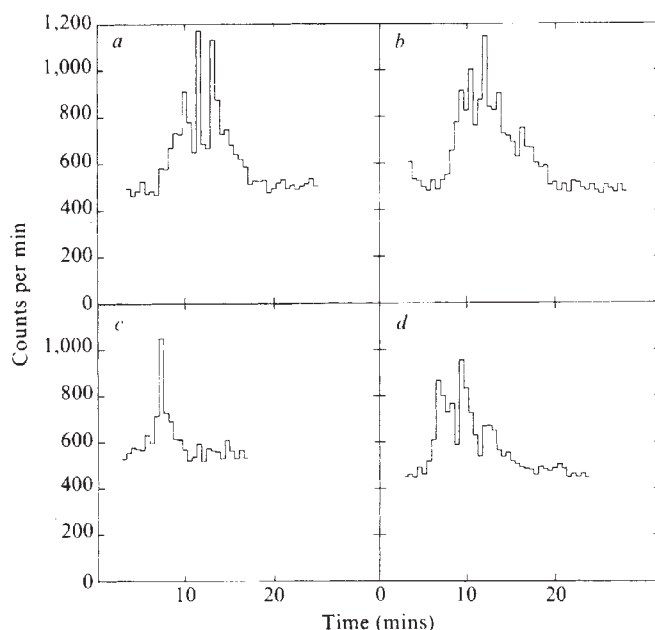


Fig. 3 Four outbursts plotted with 64 s time resolution. The times (a, b, c, d) at which they occurred are indicated in Fig. 1. The 1σ uncertainties are the square root of the count rate.

events of comparable luminosity from the vicinity of LMC X-4 and these may be related to the four outbursts reported here.

I thank G. F. Carpenter of Birmingham University and P. J. N. Davison for useful discussions.

N. E. WHITE

Mullard Space Science Laboratory,
University College London,
Holmbury St Mary,
Dorking, Surrey, UK

Received 10 October; accepted 3 November 1977.

1. Sandaleak, N. & Philip, A. G. D. *IAU Circ.* No. 3023 (1977).
2. Chevalier, C. & Ilovaisky, S. A. *Astr. Astrophys.* **59**, L9 (1977).
3. Giacconi, R. *et al.* *Astrophys. J.* **178**, 281 (1972).
4. White, N. E., Davison, P. J. N. & Carpenter, G. *IAU Circ.* No. 3100 (1977).
5. White, N. E. & Carpenter, G. *Mon. Not. R. astr. Soc.* (submitted).
6. Epstein, A. *et al.* *Astrophys. J.* **216**, 103 (1977).

A search for rapid optical periodicities from Cyg X-1

ACCORDING to the most widely accepted model of the X-ray source Cyg X-1, the X rays are produced in a gas disk surrounding and accreting into a black hole with a mass of about $10M_{\odot}$. The accretion disk is maintained by matter transferred from the binary companion of Cyg X-1, the blue supergiant HDE226868. The X-ray luminosity of Cyg X-1 is known to vary randomly on a wide range of time scales, from many days to a few milliseconds, with the shortest of these time scales corresponding to the dynamical time scale of the innermost stable orbit around the black hole^{1,2}. The optical luminosity of the Cyg X-1 system is also known to vary on time scales of days³, but only one case of extremely rapid optical variability has been reported^{4,5}. The properties of the rapid variations were most unusual. The variations were transient and were detected only four times, but they occurred sufficiently frequently that three of these detections occurred during the total of only 2 h of observations accumulated by Auriemma *et al.* in July 1975^{4,5}. The variations were nearly periodic, with periods of 83.53 ms, 83.71 ms, 83.59 ms, and with $|\Delta P/P| \sim 2 \times 10^{-4}$. The periodicities were long lived, lasting at least 6 min and probably over 10 min, so that the pulse trains consisted of over 4,000 pulses. The amplitudes of the pulses were large, up to 0.042 mag. No comparable behaviour has been detected in the X-ray light curve of Cyg X-1, and it is far from clear that this behaviour is consistent with our present conceptions of accretion disks. Thus, the detections of these pulses, if confirmed, would be of some importance since they would require a revision of the models for Cyg X-1. Therefore, we have taken an extensive set of observations of Cyg X-1 with the purpose of detecting additional examples of the rapid periodicities.

The data were acquired using one or other of the high-speed photometers mounted on the 0.76 m, 0.92 m, or 2.08 m telescopes at McDonald Observatory. Our observational technique consisted of acquiring individual 10-min long light curves of HDE-226868. The light curves were taken through a wide variety of filters, most notably the Johnson U, B, and V filters, the Strömgen v filter, and a 30 Å filter centered on the $\lambda 4,686$ Å line of He II. Each of the light curves was tested for the presence of periodicities at the time it was observed. Our technique for testing for periodicities has already been described in detail⁶. Briefly, for a sinusoidal signal, the test is equivalent to calculating the power spectrum of the light curve, but because of limitations on computer memory and speed, the spectrum is calculated only for a restricted range of frequencies. For Cyg X-1, we chose a frequency range of 11.718 to 12.216 Hz. This frequency range is centred on, and is a factor of nearly 20 greater than the range of the frequencies detected by Auriemma *et al.* We accumulated over 90 h of light curves spaced over 65 nights in September to November 1976 and in May to October 1977. With the exception of a few observations heavily contaminated by clouds, the maximum amplitude of any possible periodic signal during this

time was 0.0029 mag, which is more than a factor of 10 less than the amplitude of the variations detected by Auriemma *et al.* With no exceptions, the power at all frequencies in the tested range was always consistent with white noise.

There are several ways to account for the difference between our results and those of Auriemma *et al.* Cyg X-1 could have changed its properties between 1975 and 1976 so that the periodicities are no longer present or occur far less often. The variations could have been present during our observations, but with periods outside the range of periods we tested. It is also possible that the periodicities detected by Auriemma *et al.* were spurious and of instrumental origin.

E. L. ROBINSON
R. E. NATHER
J. AFRICANO
B. SMITH

McDonald Observatory and Department of Astronomy,
University of Texas,
Austin, Texas 78712

Received 11 October; accepted 3 November 1977.

1. Boldt, E., Holt, S., Rothschild, R. & Serlemitsos, P. *Proc. Calgary Conference on X-rays in Space* (1974).
2. Oda, M., Doi, K., Ogawara, Y., Takagishi, K. & Wada, M. *Astrophys. Space Sci.* **42**, 223 (1976).
3. Walker, E. N. *Goddard Symp. X-ray Binaries* 521 (1975).
4. Auriemma, G. *et al.* *Nature* **259**, 27 (1976).
5. Auriemma, G., Cardini, D., Costa, E., Giovannelli, F. & Ranieri, M. *Goddard Symp. X-ray Binaries* 485 (1975).
6. Nather, R. E., Robinson, E. L., Van Citters, G. W. & Hemenway, P. D. *Astrophys. J. Lett.* **211**, L125 (1977).

Quasi-periodic fluctuations in electron content during a partial solar eclipse

CHIMONAS and Hines^{1,2} have suggested that a solar eclipse might generate gravity waves in the atmosphere. During solar eclipse the localised cooling of the atmosphere in the lunar shadow causes an energy imbalance and the shadow moving at supersonic speed across the Earth's surface could be a continuous source of gravity waves propagating to great distances in the atmosphere. These quasi-periodic wave perturbations in the ionospheric electron density, caused by the coupling between the ionised and the neutral particles, have been detected by various investigators³⁻⁵ at middle and high latitude stations away from the path of the eclipse shadow. Hajkowicz⁶ has reported the observation of perturbations of quasi-periods of less than 2 min, after the October 1976 eclipse. We report here the observation of quasi-periodic fluctuations in Faraday rotation angle Ω and the 1 MHz modulation phase delay ϕ of 40 MHz transmissions from ATS-6 geosynchronous satellite recorded at Trivandrum (dip $0^{\circ}57'S$, geographical longitude $76^{\circ}57'E$) during the partial solar eclipse on 29 April 1976. ϕ directly gives columnar electron density integrated along the radio ray path from the satellite to the receiver; whereas Ω gives columnar integrated electron density only up to an altitude of 2,000 km because of the weightage by the component of the geomagnetic field along the ray path. For ATS-6 to Trivandrum ray path geometry, a change of 10° in Ω will be produced by a change of 0.47×10^{16} el. m^{-2} of electron (el.) content. In contrast to this, 10° change in ϕ will be produced by a change of 0.34×10^{15} el. m^{-2} . The accuracy of measurement of both Ω and ϕ is better than 1° .

On 29 April 1976 the solar eclipse, as seen from the 300 km altitude point on the ATS-6 to Trivandrum ray path, began at 1718h (all times are IST); maximum obscuration of about 20% of the solar disk was at 1802h; the end of the eclipse was at 1843h and the sunset time was 2003h. The values of Ω and ϕ recorded on 29 April 1976 between 1420h and 1830h were scaled at intervals of 2 min. The differences between the data and the 10-min moving averages which contain all fluctuations with quasi-periods less than 50 min are plotted against time in Fig. 1; separately for

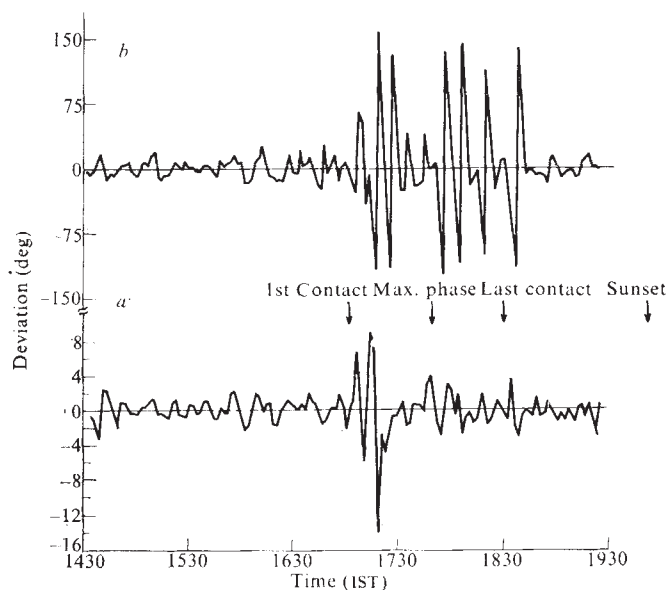


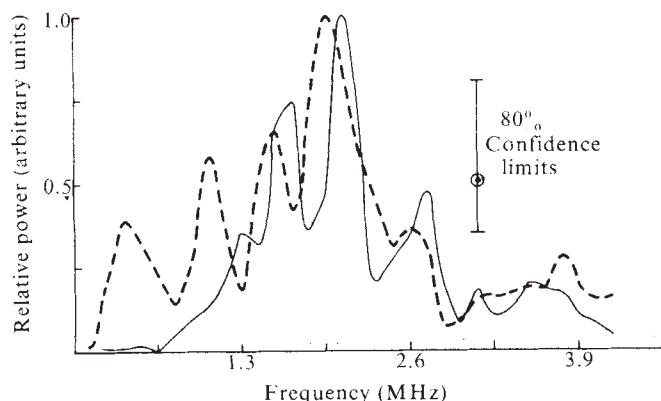
Fig. 1 Fluctuations of Ω (a) and ϕ (b) on 29 April 1976.

Ω (Fig. 1a) and ϕ (Fig. 1b). Figure 1 clearly shows fluctuations in both ϕ and Ω during the solar eclipse with quasi-periods around 11 min. A striking feature is the relatively quiet period of no fluctuations before and after the eclipse. The maximum amplitude of the total electron content fluctuations is estimated to be about 0.5×10^{16} el. m $^{-2}$ and this is about 1.0% of the total electron content. Similar analysis of the data between 1430 h and 2000 h on 28 and 30 April 1976 did not show any conspicuous quasi-periodic fluctuations. This strongly supports the view that the fluctuations shown in Fig. 1 during the solar eclipse period are in fact due to the solar eclipse.

From Fig. 1, it can be seen that there is not an exact one to one correspondence between the fluctuations of Ω and ϕ although, the ray paths along which they give the integrated electron content are the same. Also ϕ shows relatively clearer periodic fluctuations than Ω . This may be because Ω is weighted by the geomagnetic field.

To study the spectral components, the fluctuations of Ω and ϕ in Fig. 1, during the 128-min interval between 1652 h and 1900 h were subjected to power spectrum analysis with Hamming filter⁷. The power spectrum obtained has a frequency resolution of 0.13 mHz, the Nyquist frequency being 4.166 mHz. In Fig. 2 the continuous and the dashed curves are the power spectra of ϕ and Ω

Fig. 2 Power spectra of the fluctuations in Ω (dashed line) and ϕ (solid line) during the eclipse on 29 April 1976.



respectively. To test the stability of the spectral power estimates, the equivalent number of degrees of freedom of a χ^2 distribution⁷ have been computed and was found to be close to 22 for both the spectra. The 80% confidence limits of the relative spectral power estimate of 0.5 is shown in Fig. 2. The periods (in minutes) corresponding to the prominent spectral peaks in Fig. 2 are:

ϕ	—	12.8	9.8	8.0	6.1
Ω	42.7	16.0	10.7	8.5	6.5

It can be seen from atmospheric gravity wave theory⁸ that the above periods agree with the periods of gravity wave propagating modes at F-region levels. Considering the corresponding peaks in Ω and ϕ shown above it is immediately obvious that the periods are larger in Ω than in ϕ . This difference may be due to the influence of geomagnetic field on Ω .

The effect of the weightage of the geomagnetic field on Ω depends on the period and the vertical wavelength of the propagating gravity wave. For example, the electron density fluctuations at 350 km altitude produce about 50% smaller fluctuations amplitudes in Ω compared with the same electron density fluctuations at 200 km altitude. In Fig. 1, the amplitude of the fluctuations in ϕ is almost the same during the eclipse, whereas in Ω the fluctuations decreased to about 50% during the later part of the eclipse. This indicates that the fluctuations in electron density during the eclipse are not at the same height and that the fluctuations during the later part of the eclipse are at a higher altitude than soon after the eclipse beginning. A steady increase in the effective height of the electron density fluctuations in conjunction with the steady decrease in the geomagnetic field with height, leads to a larger period of the wave as seen on Ω than on ϕ . This upward energy propagation is in fact an important characteristic of the gravity waves.

In Fig. 2, the relatively larger spectral power at larger periods in Ω than in ϕ indicates that the larger period waves are propagating at lower altitudes, where for a given density change the change in Ω is larger because of the large value of the weighting geomagnetic field. This is also borne out by gravity wave theory⁸ that large period waves propagate at low elevation angles.

It can be seen that electron content fluctuations are shown better in ϕ than in Ω . The non-observation of clear fluctuations in electron content during earlier solar eclipses³⁻⁵ may be because all these measurements were of Ω fluctuations.

It is difficult to calculate the amplitude of the gravity wave produced by solar eclipse in the neutral particles and the consequent fluctuations in electron density by collisional coupling between the neutral and ionised particles. It is known⁸, however, that the gravity wave amplitude increases exponentially with height whereas the ion neutral collisional frequency decreases exponentially in the same way. This may perhaps lead to the constant amplitude of the fluctuation of ϕ in Fig. 1.

S. VAIDYANATHAN

C. RAGHAVA REDDI

B. V. KRISHNAMURTHY

Space Physics Division,
Vikram Sarabhai Space Centre,
Trivandrum-695022, India

Received 6 September; accepted 3 November 1977.

- Chimonas, G. & Hines, C. O. *J. geophys. Res.* **75**, 875 (1970).
- Chimonas, G. & Hines, C. O. *J. geophys. Res.* **76**, 7003 (1971).
- Carlson, H. C. *et al. Nature* **226**, 1124 (1970).
- Larford, G. M., Jurgens, R. M., Vesecky, J. F. & Warhann, T. W. *J. atmos. terr. Phys.* **34**, 733 (1972).
- Davies, M. & Da Rosa, A. V. *Nature* **226**, 1123 (1970).
- Hajkowicz, L. A. *Nature* **226**, 147 (1970).
- Blackman, R. B. & Tukey, J. W. *Measurement of Power Spectra* (Dover, London).
- Georges, T. M. *ESSA tech. Rep. IER 57-ITSA*, 54 (1967).

A simple calculation of ozone depletion by chlorofluoromethanes using a two-dimensional model

THE possible perturbation of the total amount of ozone in the atmosphere due to man's activities has been the subject of much recent research. The vast majority of perturbation calculations have been performed using one-dimensional models in which variations in the vertical alone are considered. I describe here some simple calculations of the effects of chlorofluoromethanes on the ozone layer using a two-dimensional (latitude and height), time-dependent numerical model in which the mean circulation is computed at each time step. In our two-dimensional treatment the eddy transports of heat and matter are calculated using two-dimensional eddy diffusion coefficients, based on atmospheric statistics, and the eddy momentum fluxes are average monthly values for 1973 deduced from the Nimbus V Selective Chopper Radiometer measurements. The approach has proved successful in modelling many features of the unperturbed atmosphere and, in particular, of the behaviour of ozone in the stratosphere¹.

In the model the photochemical source of ozone is linearised such that

$$\frac{\partial [O_3]}{\partial t} = \frac{[O_3]_e - [O_3]}{\tau}$$

where $[]$ represents a mass mixing ratio, $[O_3]_e$ is the ozone equilibrium mixing ratio and τ the time constant. $[O_3]_e$ and τ are specified as functions of latitude, height and season, the values being appropriate to a photochemical scheme including the oxides of nitrogen and hydrogen. The solar heating is dependent on the modelled ozone concentrations, thus coupling the chemistry, radiation and dynamics. Further details of the model are given in refs 1 and 2.

A complete calculation of the effect of chlorofluorocarbons on ozone would require simulation over many years. The purpose of this letter is not to predict the magnitude of these effects, however,

Fig. 1. The latitude-time section of the percentage change in total ozone between experiments E1 and E2 for one model year. Model days 730–1095.

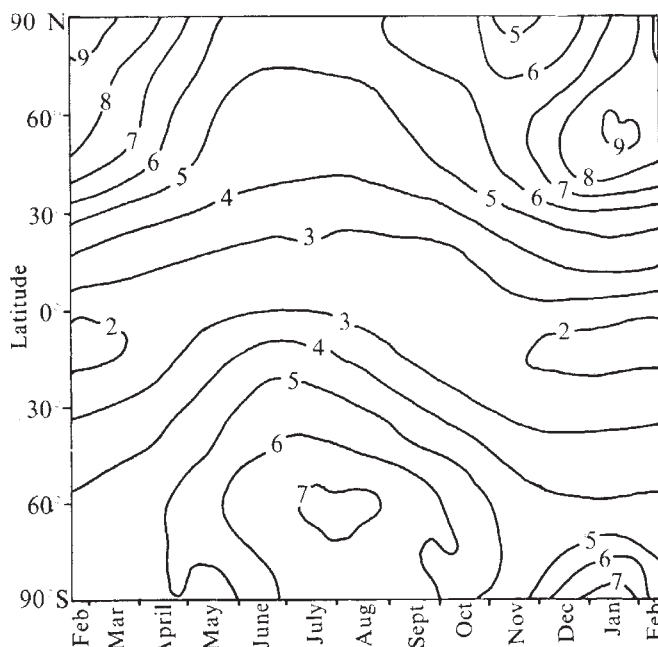
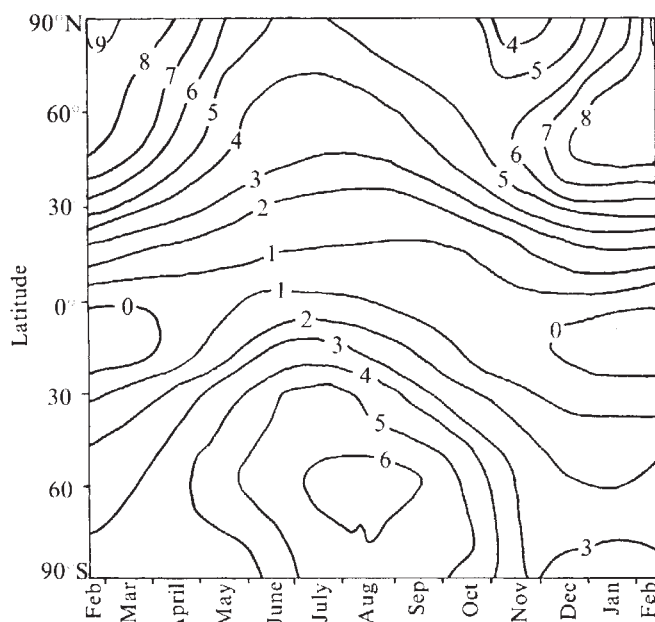


Fig. 2. The latitude-time section of the percentage change in total ozone between experiments E1 and E3 for one model year. Model days 730–1095.

but rather to indicate that there are important latitudinal and seasonal variations which need to be considered. For this reason it seems justifiable to make some rather gross approximations. The starting point for the study is the work of Crutzen³. Using a one-dimensional, steady-state model, he calculated the reduction in ozone due to continued release of chlorofluoromethanes at a rate appropriate to the mid-1970s. Large changes are predicted for the upper stratosphere where ozone is expected to be close to photochemical equilibrium.

The results of three different calculations will be described here: E1, an unperturbed run very similar to that described in Harwood and Pyle¹; E2, a perturbed run in which the equilibrium values are modified, in a manner based on Crutzen's work as described below, and which uses exactly the same temperature and wind fields as in E1; E3, an experiment in which the equilibrium values are again modified but in which the perturbed ozone concentrations are used in the calculation of solar heating. Thus, the feedback between the ozone amount, temperature and the mean circulation is included in E3.

In all three runs the ozone equilibrium concentrations are taken to be temperature dependent according to the expression $[O_3]_e(T) = A \exp(1,000/T)$ (see Barnett *et al.*⁴). The results of E1, including this temperature dependence, are similar to those described by Harwood and Pyle¹.

In the perturbed runs the equilibrium values were modified in the following manner. The stratospheric values of $[O_3]_e$ were modified by the percentage calculated by Crutzen³ (table 1 ref. 3). τ , too, depends on $[O_3]_e$ as well as on the concentration of chlorine compounds and the dissociation rate of ozone. Allowance was made for all these factors in E2 and E3.

Figure 1 shows the latitude-time section of the percentage difference in total ozone between E1 and E2. The magnitudes clearly depend on the initial assumptions and should not be emphasised. What is of interest is the large variation with latitude and time. The largest reductions occur in high latitudes where the total ozone itself is largest and at these latitudes the reductions vary by a factor of two with the time of year. In equatorial regions, where total ozone is lowest, there is even an increase in the total ozone throughout part of the year. This is due to the 'self healing' effect whereby increased penetration of solar radiation in the lower stratosphere increases ozone there. Ozone is produced in the equatorial lower stratosphere and transported in the upward branch of the Hadley cell to regions where it has been depleted.

The increased upward transport in E2 can, at certain times of the year, cancel the reductions at higher levels.

The above result is due to a combination of dynamical and photochemical processes. If a climatological ozone distribution were reduced by the same factors as $[O_3]_e$, the largest percentage reduction in total ozone would be found in equatorial latitudes. On the other hand, the percentage reduction of the equilibrium values between E1 and E2 does have largest depletion in high latitudes but in this case the magnitudes differ considerably from those in Fig. 1.

In E3 the reduced ozone concentrations are used in the solar heating calculations in the model. One might expect that reduced ozone in the upper stratosphere would lead to reduced solar heating and lower temperatures there. The ozone, though, is temperature dependent; lower temperatures cause the ozone amounts to increase. Hence, the mean circulation, driven in part by the solar heating, differs from that in E1 (and E2, where this feedback was deliberately suppressed). Figure 2 presents the latitude-time section of the percentage difference in total ozone between E1 and E3 and has many similarities with Fig. 1. Surprisingly, the reduction of ozone in E3 is larger than in E2. The stratopause temperatures in E3 are reduced by a few per cent and the ozone reduction at this level is correspondingly less than in E2. In the mid and lower stratosphere in E3, however, there is an increase in temperature and a decrease in ozone when compared with E2 and this more than cancels the upper stratospheric increase. Thus, for instance, in E3 there is no longer a small increase in total ozone at the equator. That the temperatures in the lower stratosphere are higher in E3 is probably due not only to increased solar heating but also to transport processes. It is evident from E3 that the calculation of ozone reduction is made a more complicated exercise by these important feedbacks between photochemistry, radiation and dynamics.

The results presented here demonstrate the importance of multidimensional models in studies of projected perturbations in atmospheric ozone. An extremely simple photochemical scheme has been used here and there are strong reservations about the absolute magnitude and phase of the calculated reductions in ozone. But, both E2 and E3 demonstrate, unequivocally, important variations in ozone depletion with both latitude and time (and this is confirmed by preliminary results with a model including a much more complete photochemical scheme, data not shown). Furthermore, E3 indicates the importance of including all the important feedback mechanisms.

This work was supported by the Department of the Environment.

J. A. PYLE

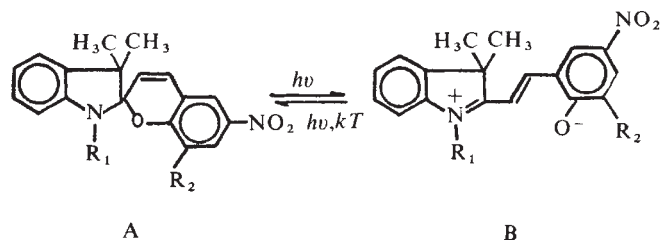
Department of Atmospheric Physics,
Clarendon Laboratory,
Oxford, UK

Received 21 September; accepted 3 November 1977.

1. Harwood, R. S. & Pyle, J. A. *Q. J. R. Met. Soc.* **103**, 319–343 (1977).
2. Harwood, R. S. & Pyle, J. A. *Q. J. R. Met. Soc.* **101**, 723–748 (1975).
3. Crutzen, P. J. *Geophys. Res. Lett.* **1**, 205–208 (1974).
4. Barnett, J. J., Houghton, J. T. & Pyle, J. A. *Q. J. R. Met. Soc.* **101**, 245–257 (1975).

Quasi-crystals from irradiated photochromic dyes in an applied electric field

In many photochromic systems, and certainly in the spiro-pyrans, the colourless (A) form is much less polar than the photocoloured (B) form¹. Our previous experiments^{2,3} revealed unusual properties of spiro-pyrans in non-polar solvents at low temperatures (173–240 K). On irradiation of these solutions with suitable light there were formed dimers (AB) and charge transfer complexes (CTC $\equiv A_n^+B_n^-$ with $n \approx 2-3$), and these in turn formed aggregates with degrees of association $>10^6$. The CTC exhibit spectral red shift of about 100 nm as compared to the dimers. The



relative populations of dimers and charge transfer complexes in the aggregates depended both on light intensity and temperature, because CTC formation involves an activation energy of about 5 Kcal mol⁻¹ associated with the interaction of dimers with further A molecules. At temperatures sufficiently low or at light intensities sufficiently high it was possible to obtain aggregates composed of dimers only. When the constant field was applied during irradiation, threads were formed which extended from one electrode to the other (along the electric lines of force). These threads are composed of dimers and CTC, and their absorption spectrum is identical to that of aggregates in solution. Linear dichroism measurements showed that the CTC are oriented along the thread axis, but the dimers are unoriented. These threads were formed only in the

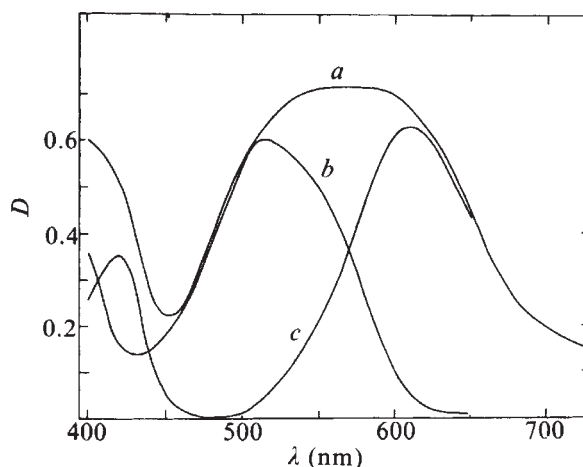


Fig. 1 The absorption spectra of 6-nitro BIPS (concentration in methylcyclohexane solution 5.10^{-4} M), irradiated by ultraviolet light with intensity of 2.10^{-8} einstein cm^{-2} s. a, Quasi-crystals, room temperatures, $F = 15 \text{ kv cm}^{-1}$; b, aggregates -120°C . c, Difference spectrum (absorption of the CTC).

temperature range in which CTC arise in solution. Here we describe investigations of the structure of such threads, which led us to conclude that the threads consist of highly dipolar crystallites composed of the charge transfer complexes and coated by an amorphous phase composed of dimers. These crystallites in the amorphous envelope are oriented along the electric field and joined together to give the quasi-crystalline threads.

Solutions of 6-nitro-1',3',3'-trimethylspiro-/2H-1-benzopyran-2,2'-indoline, '6-nitro BIPS' ($R_1 = \text{CH}_3$; $R_2 = \text{H}$) (ref. 4), in methylcyclohexane were studied. An Osram HBO 200 W high pressure mercury lamp in a thin housing was used for irradiation, with Corning glass filter No. 5840 which transmits essentially from 300 to 400 nm. The absorption spectra were measured using a Cary 14 spectrophotometer, and the linear polarised spectra using the same instrument with the attachment described by Jaffe *et al.*⁵. This attachment permits simultaneous measurement of absorptions in two mutually perpendicular directions. This proved useful in the present experiments because the

Table 1 Degree of linear dichroism of the threads from 6-nitro BIPS in methylcyclohexane

Concentration of solution $M \times 10^3$	Applied field $kV\ cm^{-1}$	P_{510} (± 0.03)	P_{610} (± 0.02)
3.0	5.0	0.03	0.11
3.0	7.5	0.02	0.09
1.5	25	0.00	0.1
0.8	25	-0.03	0.11
0.5	25	-0.03	0.12
0.4	25	-0.01	0.12

$P_\lambda = (D_{\parallel} - D_{\perp}) / (D_{\parallel} + D_{\perp})$; $D_{\parallel}$ and $D_{\perp}$ are the optical densities of the threads when the monitoring light is polarised, respectively parallel and perpendicular to the axes of the threads.

threads are rather unstable. But the method limited us to measurements of absorption in specific directions only, that is parallel and perpendicular to the threads axis.

We used spiropyran solutions of higher concentrations ($>10^{-4} M$) and higher light intensities ($>10^{-8}$ einstein $cm^{-2} s^{-1}$) than in the earlier studies^{2,3}. This permits the production of much higher stationary concentrations of form B and the growth of the quasi-crystals even at room temperature.

Figure 1 shows the spectrum of quasi-crystals obtained from 6-nitro-BIPS at room temperature in an electric field. This spectrum is identical to that of aggregates obtained under the same conditions in the absence of a field. Also the spectrum of aggregates obtained at 153 K in solution is shown, which consist of dimers only without CTC. The difference spectrum, constructed as described previously³, is the absorption due to CTC; this is shown in Fig. 1c. This absorption is red-shifted by about 100 nm with respect to the absorption of the dimers.

The linear dichroism, P_λ (see Table 1) was measured for various initial concentrations of the solution and for various strengths of the applied field. The results show that P_{510} , determined near the maximum dimer absorption, is close to zero; on the other hand, P_{610} , determined near the maximum CTC absorption, is equal to about 0.1, and does not vary with the concentration of the photochrome or the field strength (5–25 $kV\ cm^{-1}$).

We shall show that the threads consist of small crystallites of CTC embedded in dimers. The dichroism results then show that there is a random arrangement of the dimers in the threads; on the other hand, the microcrystals in the thread are all aligned, and in such a way that the absorption by the CTC of the light polarised parallel to the thread axis exceeds, although slightly, the absorption of the light polarised perpendicular to this axis. This orientation of the microcrystals must be due to the interaction $F\mu$ between the field of strength F and some permanent dipole moment μ , with $F\mu > kT$. This inequality is satisfied at 293 K and for $F = 5\ kV\ cm^{-1}$ when $\mu > 2.5 \times 10^{-30}$ D. This corresponds to a dipole length $>10^{-5}$ cm which considerably exceeds the size of the complexes $A^+ \cdots B^-$ ($n = 2-3$). The conclusion is that it is associates or microcrystals of CTC rather than the individual CTC which are oriented by the applied field.

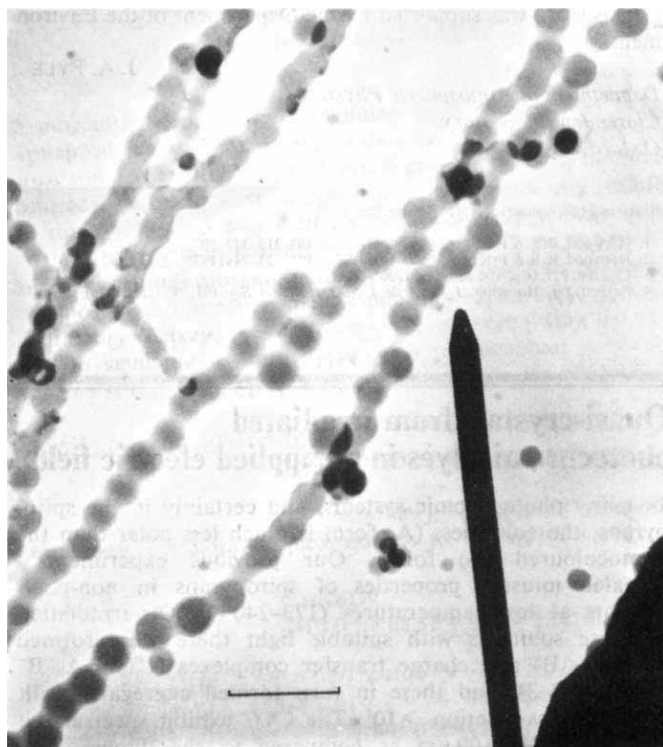
We recall that the photocoloured B form of the spiropyrans are merocyanines which readily aggregate in solution to give stacks with a tilted 'deck-of-cards' structure⁴. The absorption of such aggregates are red-shifted by about 100 nm as compared to the absorptions of the corresponding unaggregated molecules. In our case, the stacks as well as the crystals which are formed from the highly dipolar CTC may or may not be dipolar depending on their structures. Our evidence points to the crystals being oriented by the field. Thus we conclude that the crystals are dipolar. Since our measurements of polarised absorption were limited to directions parallel and perpendicular to the thread axis, we

cannot tell whether the observed value of P_{610} is its maximum value, and cannot ascertain the relationship between the orientations of the CTC chromophore and the crystal dipole moment.

Additional structural information has been obtained by X-ray and electron diffraction studies. The X-ray powder patterns of precipitates formed at room temperature in the absence of an applied field and of the quasi-crystalline threads are practically identical; they consist of sharp reflections, the most intense ones corresponding to interplanar spacings of 3.4, 5.0, and 7.4 Å. These sharp reflections are superimposed on a diffuse pattern, the intensity of which rises with increasing dimer content. Thus precipitates formed at low temperature (150 K), which are composed of dimers only, with no CTC, give only the diffuse pattern. Therefore both the precipitates and the quasi-crystals are composed of various proportions of dimers, in a more-or-less amorphous state, and of crystalline CTC.

Electron microscope studies of precipitates formed in the absence of a field show them to consist of globules, $\sim 0.2-0.4\ \mu m$ in thickness. For the electron microscope studies of the threads the latter were produced directly on a parlodion film, carefully dried, and transferred to grids. In the micrograph (Fig. 2) one can see similar globules of those of the precipitates, joined together by straight, narrow portions of the threads, which also prove to be amorphous. Because of the thickness of the globules it is not possible to get distinct electron diffraction patterns from their central parts, but such patterns can be obtained sometimes by focussing on the edges. The patterns thus obtained consist of the reflections from one or more single crystals. The diffraction from the crystals can be readily detected in those cases in which the amorphous envelopes of the globules have been (accidentally) removed. We conclude that crystals are contained in the globules, and that their dimensions are of the same order of magnitude as estimated from spectroscopic data. Further, since limiting dichroism is achieved even at the low field-strength of $5\ kV\ cm^{-1}$, the microcrystals seem to be aligned parallel in the threads.

During irradiation of spiropyrans, in non-polar solvents

Fig. 2 Electron micrograph of the quasi-crystals. Scalebar, $1\ \mu m$.

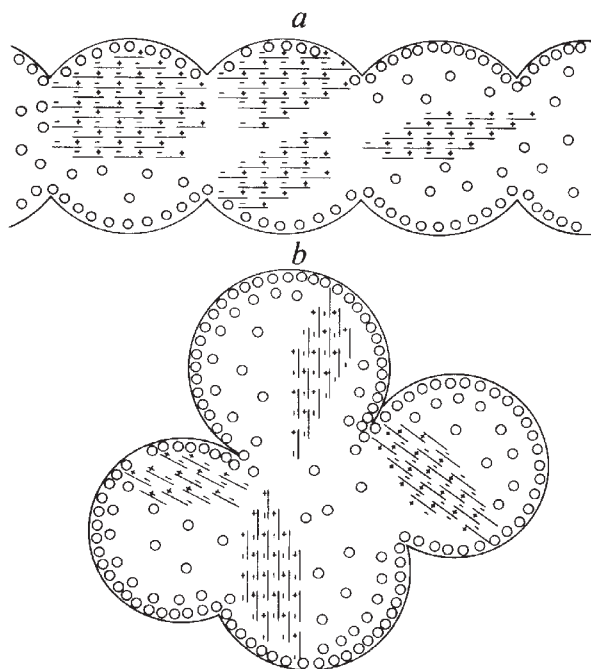


Fig. 3 Scheme of structures of quasi-crystals (a) and aggregates (b).

dimer and charge-transfer complexes are formed. These give solids consisting of amorphous zones of dimer and dipolar crystals of the complexes. In an electric field the crystals coated with dimer are joined together to give quasi-crystalline threads in which the crystals are oriented with their dipolar axes along the lines of force and thread axes. These threads thus form a new, linearly organised form of organic material (Fig. 3).

We thank Professors M. D. Cohen and E. Fischer, and Dr E. Wachtel, for valuable discussions, and Mr M. Kaganowich for the spiropyran.

V. A. KRONGAUZ

E. S. GOLDBURT

Department of Structural Chemistry,
The Weizmann Institute of Science,
Rehovot, Israel

Received 23 August; accepted 3 November 1977.

1. Bertelson, R. *Photochromism* ch. 3 (ed. Brown, G. H., Wiley, New York, 1971).
2. Krongauz, V. A. & Parshutkin, A. A. *Photochem. Photobiol.* **15**, 503 (1972).
3. Parshutkin, A. A. & Krongauz, V. A. *Molec. Photochem.* **6**, 437 (1974).
4. Koelsch, C. F. & Workman, W. R. *J. Am. chem. Soc.* **74**, 6288 (1952).
5. Jaffe, J. H., Jaffe, H. & Rosenhek, R. *Rev. Scient. Instrum.* **38**, 935 (1967).
6. Mees, C. E. & James, T. H. *Theory of the Photographic Process* ch. 11 & 12 (Macmillan, New York, 1966).

Rare gas isotopic compositions in diamonds

RARE gas isotopic compositions such as $^{40}\text{Ar}/^{36}\text{Ar}$, $^3\text{He}/^4\text{He}$ and $^{129}\text{Xe}/^{132}\text{Xe}$ in the Earth have provided a powerful tool for understanding the origin and evolution of the terrestrial atmosphere¹⁻⁴. The isotopic information may be obtained from rare gases trapped in some mantle-derived materials such as volcanic rocks, volcanic xenoliths or volcanic gases. Among these mantle-derived materials, diamond seems to be unique due to its almost complete inertness to any known chemical and to its enormous stability against high temperature. Although the presence of O_2 , H_2 , CH_4 , H_2O , CO , N_2 , Ar and CO_2 in diamonds has been

reported^{5,6}. No previous measurement has been made either on elemental compositions or on isotopic ratios of rare gases in diamonds. Here we report on rare gas elemental composition and isotopic ratios in diamond. We found that $^3\text{He}/^4\text{He}$ ratio is more than an order of magnitude larger than the atmospheric value and also $^{40}\text{Ar}/^{36}\text{Ar}$ ratio is significantly higher.

Industrial diamonds we studied are believed to have come from Kimberley Mines, South Africa. The size of the diamonds ranges from about 1 mm to about 5 mm. Some contain black inclusions, some of which are ferromagnetic and appears to be pyrrhotite under a microscope. Existence of pyrrhotite inclusions in diamond is not unusual⁷ and the black inclusions seem to be syngenetic with diamonds. Diamonds are crushed to a few meshes in a stainless steel mortar. Step-heating was made on two batches of samples at 800 °C, 2,000 °C and 2,100 °C in a high vacuum tantalum furnace⁸. Only visible difference between the two batches is that batch 1 is richer in black inclusions than batch 2.

Rare gas elemental abundance and their isotopic ratios are shown in Table 1. Batch 1 gave almost three times more rare gases as batch 2. Also isotopic ratios, $^{40}\text{Ar}/^{36}\text{Ar}$ and $^3\text{He}/^4\text{He}$, of the two batches are quite different. We believe that this is due mainly to difference in the amount of the black inclusions, that is, rare gases in batch 1 may come mostly from the black inclusions, while those in batch 2 are more representative of diamonds. In both cases major degassing is observed at 2,000 °C. A third batch of sample which were heated at 1,500 °C gave off $^{40}\text{Ar} \sim 1 \times 10^{-8} \text{ cm}^3 \text{ g}^{-1}$ at STP which is less than 10% of the rare gases evolved from batch 1 and 2 at 2,000 °C. X-ray analyses on heated samples showed that heating at 1,500 °C for 1 h (batch 3) did not result in graphitisation, but heating at 2,000 °C for 1 h (batch 1 and batch 2) completely converted diamonds to graphite. We conclude, therefore, that graphitisation which accompanies any density change of more than 30% must be responsible for the major degassing of diamonds.

All other rare gases except He, show a similar thermal release pattern. In the case of He, more than 99% was degassed at 2,000 °C suggesting that most of He resided at low temperature sites had been degassed before the experiment, perhaps under high temperature mantle condition. From the similar thermal release pattern of rare gases between batch 1 and batch 2, we conclude that rare gases in the black inclusions occupy similar sites as those in batch 2, the latter being more likely to represent diamond rare gases. This suggests a syngenetic origin of the black inclusions with diamonds.

The K content of evaporates which were deposited on the inner wall of a high vacuum extraction furnace during the graphitisation were measured. The deposits were washed by dilute HNO_3 and the K content in the washed solution was measured with an isotope dilution method. The amount of potassium was 8.22 p.p.m. and 1.97 p.p.m. for batch 1 and batch 2 respectively. Radiogenic ^{40}Ar , which would have been produced in batch 1 diamonds by this amount of K, will exceed the total amount of ^{40}Ar if the age of the diamonds is older than 2,800 Myr. Hence, the diamonds can not be older than 2,800 Myr. If this maximum age is assumed, this will give minimum $^{40}\text{Ar}/^{36}\text{Ar}$ ratio of 670 for the trapped Ar in the batch 2 diamonds. The observed $^3\text{He}/^4\text{He}$ ratio is also minimum, as the diamonds must contain some U as well as K. Therefore, the $^3\text{He}/^4\text{He}$ and $^{40}\text{Ar}/^{36}\text{Ar}$ in the diamonds must be larger than 1.95×10^{-5} and 670. The $^3\text{He}/^4\text{He}$ ratio is significantly higher than those observed in ridge submarine glasses⁹ which are generally regarded to be derived from oceanic upper mantle. The high $^3\text{He}/^4\text{He}$ ratio suggests that the diamonds may be derived from a region different from source regions for oceanic ridge basalts, most likely from deeper region in the mantle. In this connection it is interesting that similar high $^3\text{He}/^4\text{He}$ (2.09×10^{-5}) are observed in Hawaiian volcanic gases⁹, which may be derived from deeper mantle through a hot spot.

Although the origin of diamonds is not fully understood, it is generally assumed that they crystallise from CO_2 saturated fluid phase in the mantle, probably deeper than 100 km or even more. The syngenetic black inclusions occluded in diamonds seem to have been derived from mantle region in which diamonds crystallised. Our results lead us to an interesting speculation that

Table 1 Elemental and isotopic compositions of rare gases in diamond

Sample	Batch 1			Blank 1		Batch 2		Blank 2
Weight (g)		1.98				2.61		
Temperature (°C)	800	2,000	2,100	2,000	800	2,000	2,100	2,000
Time of heating (min)	30	60	5	60	30	60	60	60
$^3\text{He} \times 10^{-13} \text{ cm}^3 \text{ g}^{-1} \text{ (STP)}$	<3	286	<10	<20	<10	179	<10	<20
$^4\text{He} \times 10^{-9}$	1.36	3480	0.09	1.1	1.57	918	0.37	2
$^{20}\text{Ne} \times 10^{-12}$	6.3	20.7	<0.1	20	2.1	10.2	<0.1	20
$^{36}\text{Ar} \times 10^{-11}$	3.35	36.7	10	2.4	1.54	9.37	1.90	2
$^{84}\text{Kr} \times 10^{-12}$	1.8	3.6	<0.02	0.6	1.0	2.9	0.26	0.8
$^{132}\text{Xe} \times 10^{-13}$	3.5	9.1	<2	2	4.4	7.3	1.0	2
$^3\text{He}/^4\text{He}$	$<2 \times 10^{-4}$	$(8.23 \pm 0.35) \times 10^{-6}$				$(1.95 \pm 0.07) \times 10^{-5}$		
$^{40}\text{Ar}/^{36}\text{Ar}$	359 ± 2	436 ± 2		574 ± 14		$1,121 \pm 8$		
$^{128}\text{Xe}/^{132}\text{Xe}$		0.0754 ± 0.0011				0.0718 ± 0.0037		
$^{129}\text{Xe}/^{132}\text{Xe}$		0.996 ± 0.019				0.978 ± 0.004		
$^{130}\text{Xe}/^{132}\text{Xe}$		0.161 ± 0.007				0.149 ± 0.006		
$^{131}\text{Xe}/^{132}\text{Xe}$		0.786 ± 0.023				0.779 ± 0.014		
$^{134}\text{Xe}/^{132}\text{Xe}$		0.391 ± 0.012				0.386 ± 0.009		
$^{136}\text{Xe}/^{132}\text{Xe}$		0.328 ± 0.012				0.318 ± 0.018		

Other isotopic ratios are indistinguishable from atmospheric values

rare gases observed in batch 2 may be representative of those carried by CO_2 phase, whereas rare gases in batch 1, or more specifically rare gases in the black inclusions, is rather indicative of a region where the diamonds occluded the former.

Ar isotopic ratios observed in diamonds do not seem to support the speculation¹⁰ that diamonds were formed from materials in subducting ocean sediments, since even very small amount of seawater contamination (say less than 0.1%) which would necessarily accompany ocean sediments would reduce $^{40}\text{Ar}/^{36}\text{Ar}$ isotopic ratio to almost atmospheric value (295.5) because of relatively very high Ar content (about 0.03%) in seawater. The observed $^{40}\text{Ar}/^{36}\text{Ar}$ is much higher than the atmospheric ratio, ruling out any noticeable amount of seawater contamination.

NOBUO TAKAOKA

Department of Physics,
Osaka University,
Toyonaka 560, Japan

MINORU OZIMA

Geophysical Institute,
University of Tokyo,
Tokyo 113, Japan

Received 2 August; accepted 4 November 1977.

1. Ozima, M. *Geochim. cosmochim. Acta* **39**, 1127-1134 (1975).
2. Schwarzman, D. W. *Nature Phys. Sci.* **245**, 20 (1973).
3. Tolstikhin, I. N. *Earth planet. Sci. Lett.* **26**, 88-96 (1975).
4. Boulos, M. S. & Manuel, O. K. *Science* **174**, 1334-1336 (1971).
5. Melton, C. E. & Giardini, A. A. *Am. Mineral.* **59**, 775-782 (1974).
6. Melton, C. E. & Giardini, A. A. *Nature* **263**, 309-310 (1976).
7. Sharp, W. E. *Nature* **211**, 402-403 (1966).
8. Takaoka, N. *Mass Spectrosc.* **24**, 73-86 (1976).
9. Craig, H. & Lupton, J. L. *Earth planet. Sci. Lett.* **31**, 369-385 (1976).
10. Frank, F. C. in *Sci. Technol. Ind. Diamonds* **1** (1967).

On melting icebergs

THE feasibility of towing icebergs to their coasts and melting them on arrival to provide a supply of fresh water is being studied in various parts of the world. Saudi Arabia, Australia and California¹ are amongst those considering such a project. One suggestion for melting the icebergs is to run each iceberg aground in water approximately 250 m deep, the mean depth of Antarctic icebergs. A relatively shallow pen would then be built around the iceberg and it is conjectured that the melt water will rise, without much mixing, into the pen, from where it will be siphoned off for subsequent use. Neshyba² claims, however, that melt water produced by icebergs may be responsible for a large amount of upwelling and mixing in the Antarctic's Weddell Sea. He asserts that as the melt water rises up the side of an iceberg, it entrains a sizeable quantity of warmer, saltier water from

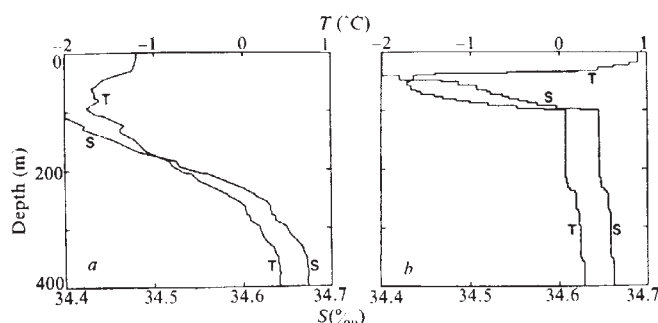


Fig. 1 Temperature (T) and salinity (S) profiles measured in the upper 400 m of the Weddell Sea (adapted from Fig. 3 of Foster and Carmack³). *a*, Near the Scotia Ridge at the northern edge of the Sea; *b*, near the turning point of the current gyre.

the environment. Neshyba calculates that an average-sized iceberg is thus responsible for a vertical volume transport into the upper layer of the Weddell Sea of $3.3 \times 10^6 \text{ cm}^3 \text{ s}^{-1}$. This is a considerable transport, and the raising of water and nutrients from depth by this mechanism would be of importance in determining the physical and biological properties of the Weddell Sea. The first suggestion, that fresh melt water rises to the surface, and Neshyba's claims that the melt water will have mixed significantly with its salty environment before reaching the surface, cannot both be correct. We show here that both may be in error.

The vital fact neglected in both arguments is that the ocean environment is not only salty, but there is also a salinity gradient in the upper 250 m of most areas of the oceans. For example, data from the Weddell Sea shown in Fig. 1 (taken from Foster and Carmack³) indicate that the salinity increases with depth, as does the temperature in this region. The data indicate further that both the salinity and temperature distribution have a characteristic 'stepped' structure, with layers of well mixed temperature and salinity separated by sharp interfaces in which both properties vary more rapidly. Similarly, off the Californian coast the salinity again has a general increase with depth, while there is an overall decrease in temperature (see, for example, Gregg and Cox⁴, in particular their Fig. 1). We have carried out a series of experiments in which ice is inserted into a salinity gradient, and the preliminary results are reported here.

Iceblocks made from water which had been deaerated under vacuum measuring approximately $20 \text{ cm} \times 10 \text{ cm} \times 3 \text{ cm}$ were inserted vertically (see Figs 2-3) into salt-stratified water at room temperature ($\sim 23^\circ \text{C}$) with constant density gradients of between 1×10^{-3} and $6 \times 10^{-3} \text{ g}$

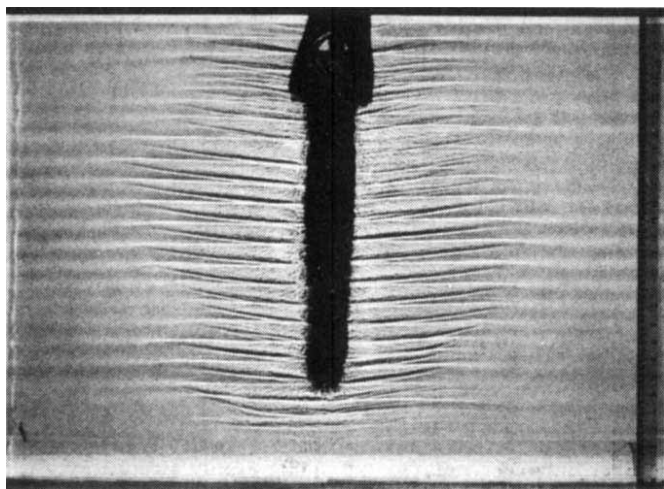


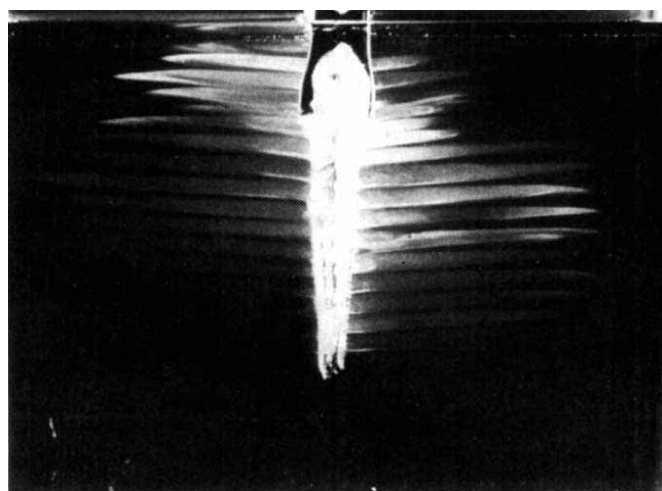
Fig. 2 Shadowgraph photograph showing the tilted layers and interfaces produced by inserting a block of ice into salt-stratified water at room temperature. (Specific gravity at top of tank 1.00, and at bottom 1.05; depth 26 cm.)

cm^{-3} per cm and surface densities between 1.00 and 1.10 g cm^{-3} .

Adjacent to the ice, melt water at 0°C rises in a thin, generally turbulent boundary layer until its density equals that of the surroundings at the same level. Above the level where these densities are equal the melt water in the boundary layer sinks. In our experiments this level occurs at a density of approximately 1.01 g cm^{-3} (if this exists), although in the ocean, with its typically lower temperature and higher salinity, such a level (where the density of the fresh melt water equals that of its surroundings) is unlikely to occur. The presence of a horizontal temperature gradient imposed upon the existing vertical salinity gradient alters the boundary layer a small distance away from the ice, producing a regular series of tilted convecting layers which grow out into the environment, as shown in Figs 2–4.

These layers exist because of double-diffusive convection, which readily occurs whenever two components of different molecular diffusivities contribute to the density. The general principles governing this form of convection and some applications are discussed by Turner⁵. Along the bottom of each layer, relatively fresh, cold water acquires heat and some

Fig. 3 An experiment carried out using the same conditions as shown in Fig. 2, but with fluorescein frozen into the ice block, and side illumination. The spread of the dye again shows the layers, but also indicates the distribution of the melt water.



salt across the 'diffusive' interface separating it from the environmental water at the top of the layer below. Thus the melt water becomes progressively lighter and runs uphill as indicated by the upward tilt of the layers seen in Figs 2–3. Along the top of the layers there is an inflow of environmental fluid, which is directly mixed with the melt water only near the ice. In this way a considerable portion of the melt water can be fed into the environment close to the level at which it is produced and very little, if any, rises to the surface.

The melting rate of the ice is controlled by the rate of heat transfer from the environment, and is thus greatly influenced by the circulation that occurs. On a large scale this means that increasing the salinity of the environmental fluid increases the relative buoyancy of the melt water.

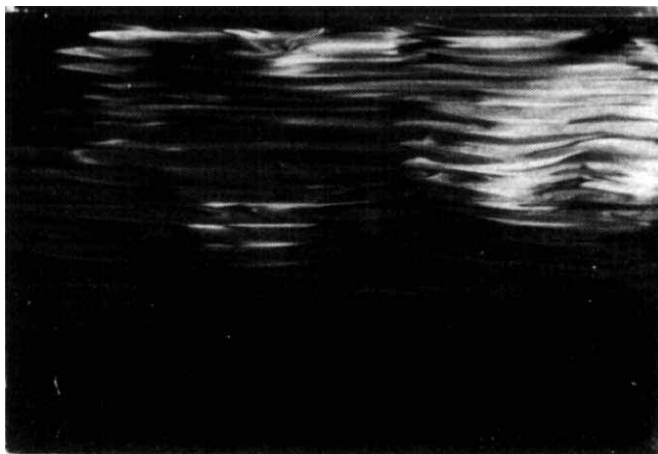


Fig. 4 The experimental tank at a later stage of the experiment of Fig. 3, soon after the iceblock had been carefully removed. The volume of dyed water is much greater than that of the ice block, indicating that the melt water mixes with a large volume of salty water and spreads out close to the level at which it is produced; little reaches the surface.

The velocity in the boundary layer thus increases, which results in increased melting. On a smaller scale, the layered motion produces a series of horizontal ridges (see Fig. 5) along the sides of the ice, with the deepest erosion corresponding to the centre of each of the layers, and the ridges to the interfaces between them. In each experiment, the layer size is reasonably uniform with depth and increases with either decreasing density gradient or increasing central salinity. In the laboratory this size is of the order of 1 cm. In the Antarctic, the weaker salinity gradients will tend to make the layers deeper, though this will be counteracted by the smaller horizontal density gradients. The quantitative scaling relations are to be discussed elsewhere.

Further suggestions follow immediately from the results of these experiments. First, the data displayed in Fig. 1b were obtained near the centre of the large gyre which characterises the circulation in the Weddell Sea. This structure differs markedly from that at the edges of the gyre (Fig. 1a) in that it has much stronger temperature and salinity gradients from 50–100m depth, weaker density gradients from 50–500 m and a more clearly defined step structure. We conjecture that this difference is due to the systematic injection of melt water at intermediate depths by icebergs which spend a longer time near the centre of the gyre. The spatial variability of the layering in this region³ is consistent with such a localised mechanism of formation, rather than a more extensive 'one-dimensional' convection process acting over the whole area.

Further measurements of temperature and salinity, made as close as possible to an iceberg in order to detect the layering, would be instructive. It would then be worthwhile

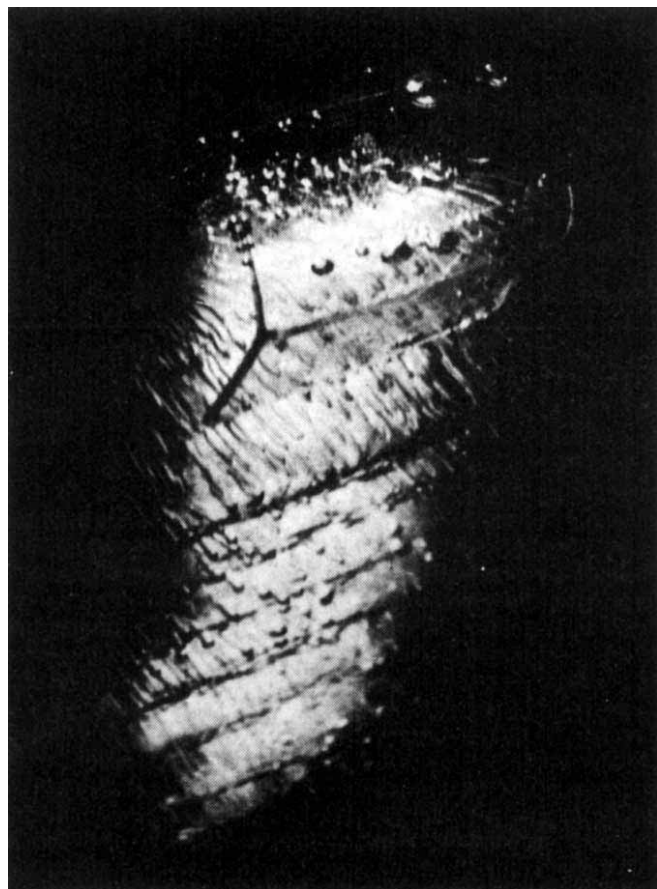


Fig. 5 An oblique view of a melting ice block, taken through the water surface. Note the ridges, caused by the uneven melting associated with the convection in layers.

exploring the structure of the layers as a function of distance from the ice. A direct search for the ridges associated with the variation in rate of melting produced by the convection might also be considered.

Finally, our results imply that those seeking to obtain fresh water from icebergs will have to consider alternative methods which isolate the melt water in some way from the surrounding seawater.

We thank R. Wyld-Browne for his able assistance with the experiments and particularly with the photography. This research was carried out while H.E.H. was a visiting fellow at the Australian National University and was in part supported by grants from the Royal Society and the British Admiralty.

HERBERT E. HUPPERT

Department of Applied Mathematics
and Theoretical Physics,
Silver Street,
Cambridge, UK

J. STEWART TURNER

Research School of Earth Sciences,
Australian National University,
Box 4, Canberra, Australia

Received 21 September; accepted 3 November 1977.

1. Agarwal, A. *New Scientist* 75, 11–13 (1977).
2. Neshyba, S. *Nature* 267, 507–508 (1977).
3. Foster, T. D. & Carmack, E. C. *J. phys. Oceanogr.* 6, 36–44 (1976).
4. Gregg, M. C. & Cox, C. S. *Deep-Sea Res.* 19, 355–376 (1972).
5. Turner, J. S. *Bouyancy Effects in Fluids* (Cambridge, 1973); *A. Rev. fluid Mech.* 6, 37–56 (1974).

Iceberg sounding by impulse radar

KNOWLEDGE of an iceberg's draft is essential for assessing its risk to underwater installations, in predicting its drift, and for estimating its total bulk. Because of the highly irregular shape of icebergs, it is impossible to estimate an iceberg's draft directly from its above-water dimensions. Large tabular icebergs have been sounded using radio techniques^{1,2}. We report here that estimates of the draft of irregularly-shaped icebergs can also be obtained from the air quickly and accurately using short-pulse radar. A small iceberg in Twillingate Harbour, Newfoundland (49° 40'N, 54° 46'W) was sounded from a helicopter using impulse radar, on 11 June 1977. The result was verified by simultaneous measurement of the iceberg's draft using side-scan sonar.

The impulse radar equipment¹, built by Geophysical Survey Systems consists of a control unit and an FM tape recorder, which were mounted in a helicopter, and a transmitter-receiver and antenna assembly, which was slung in a net about 6 m beneath the aircraft. The radiation pattern of the antenna is similar to that of a half-wave dipole. The transmitted signal is a broad-band pulse with a duration of about 20 ns and a centre frequency near 80 MHz. The receiver was set to accept echoes for up to 1.4 μ s, and by sequential sampling of the received signal, an audio-frequency replica trace was constructed and recorded on tape. After sampling, the effective repetition rate was 51.2 scans s⁻¹.

One of the impulse radar records is shown in Fig. 1, taken with the aircraft flying approximately 20 m above sea level. Strong echoes are represented by dark lines; signals received with long delay time are enhanced by a time-gain amplifier. Initial returns are due to reflections of the transmitted pulse between the antenna and the helicopter. The echo from the sea surface is very clear, but, as the iceberg is approached, this echo fades and a faint return from the sharply-peaked top of the iceberg is seen. Reflections from within the iceberg are very clear, although there is no unique bottom echo, since the radar signal reverberates within the iceberg. A set of multiple echoes is received with longer delay time.

To interpret the iceberg draft, the speed of the radar pulse in ice must be known. The speed of electromagnetic waves travelling in a dielectric medium is given by $c/\epsilon_r^{1/2}$, where c is the speed of light in free space, and ϵ_r is the relative dielectric constant of the medium. A dielectric constant of 3.2 was used, giving a velocity of 0.084 m ns⁻¹ of two-way travel time in ice. Ice samples collected from the iceberg surface were polycrystalline, and had a density of 0.938 ± 0.011 Mg m⁻³. These values for density and dielectric constant are consistent with previous work³.

The travel time used to interpret the draft was that from the first subsurface echo that was received directly beneath the centre of the iceberg. This return represents the most direct route through the ice, and it had a consistent multiple. Of course, there is no guarantee that this return came from the deepest subsurface point of the iceberg. The average draft of the iceberg was estimated to be 18.0 ± 0.9 m, with a standard deviation of 0.7 m over 16 passes. The height of the iceberg was also estimated, giving a draft: height ratio of about 4.3:1.

The hyperbolic patterns seen on either side of the iceberg are interpreted as reflections from pulses which entered the upper side of the iceberg, and traversed the iceberg before returning. Their path length in the ice was greater than the direct top-to-bottom path length, and the hyperbolic pattern is caused by the increased path length in the air as the horizontal helicopter-iceberg separation increased. These hyperbolic returns are often stronger than the direct bottom return as a larger proportion of the signal incident on the

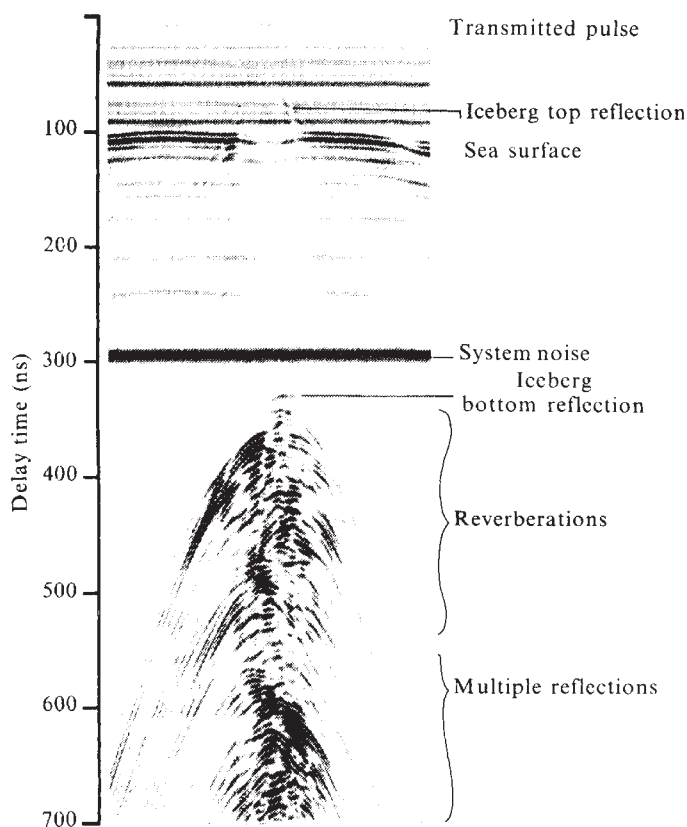


Fig. 1 Impulse radar record taken from a helicopter flying over the iceberg. Horizontal axis is aircraft position; vertical axis is radar echo delay time. Aircraft flew approximately southward during this run.

iceberg entered the iceberg through its flat upper surface than through its sharply peaked top. The hyperbolic returns contain information about the subsurface shape of the iceberg, and this might be better delineated by suitable processing of the data^{4,5}.

An estimate of the attenuation of the radar pulse within the ice was made by comparing the amplitude of the iceberg bottom echo with the amplitude of reflections from the calm sea surface, assuming a normally incident signal on plane interfaces. This assumption results in an upper estimate for attenuation. Using this approach, the attenuation rate (exclusive of spherical spreading) was estimated to be $0.20 \pm 0.02 \text{ dB m}^{-1}$, implying the possibility of sounding icebergs several hundred metres thick. The attenuation rate could easily be an order of magnitude lower, which would be more consistent with previous work on polar ice with few impurities^{2,3,6}. The chlorinity of ice samples taken near the surface of the iceberg was $1\text{--}10 \times 10^{-5}$, indicating little contamination of the ice by sea water. No evidence was

seen for random internal scattering due to inhomogeneities, common to radar sounding measurements of temperate glaciers⁷.

Side-scan sonar has been used previously to profile icebergs⁸, by lowering a sonar transducer vertically, at a known rate, from the side of a boat. The sonar system used was built by Klein Associates, and was modified for vertical operation. A model 8A-350 acoustic transceiver, which operates at 100 kHz, and a model 402 towfish, which transmits a $1.75^\circ \times 20^\circ$ fan-shaped beam, were used.

Fig. 3 Iceberg above-water form as it rolled, viewed from the south-west. Lower photograph shows approximate final position of the iceberg during the measurements. This sequence spanned about 3 min.

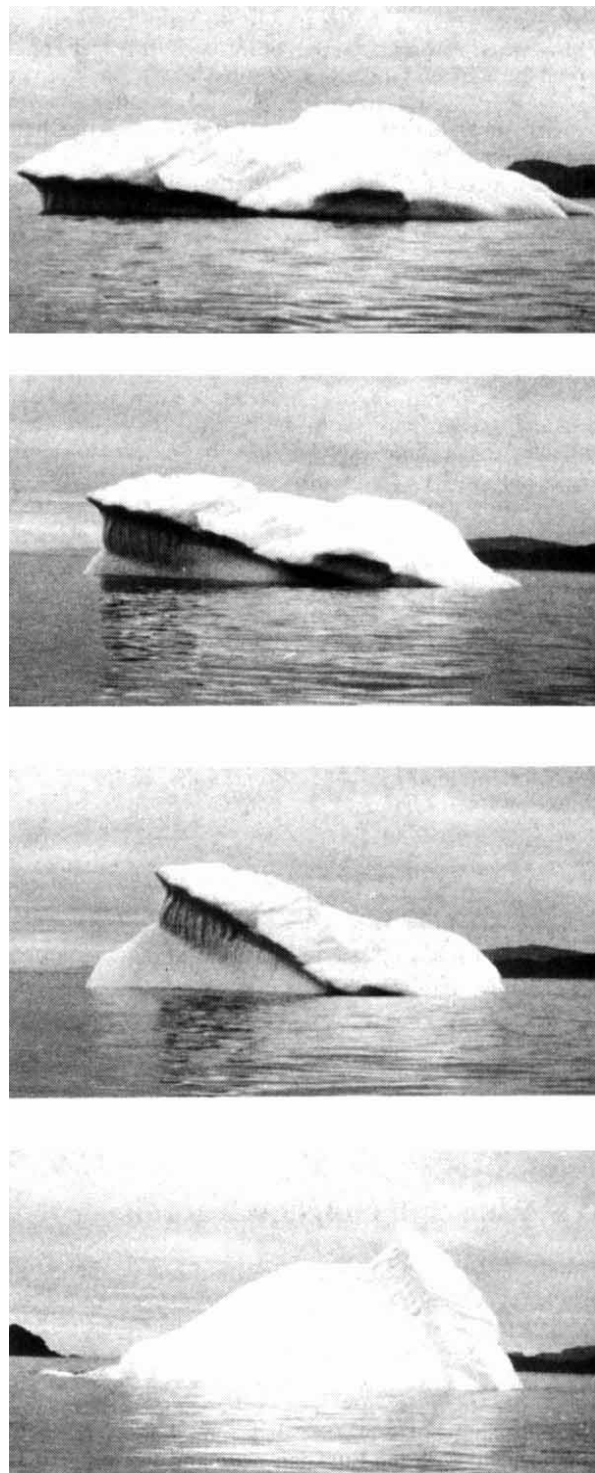
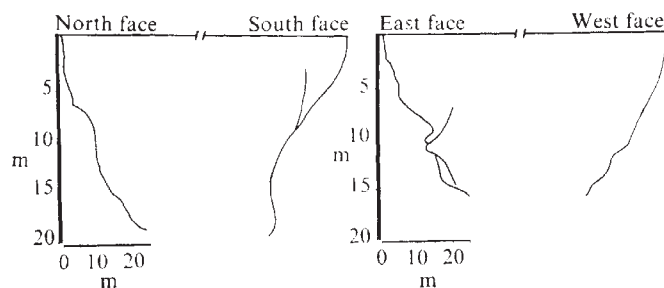


Fig. 2 Interpretation, obtained from side-scan sonographs, of iceberg underwater profile, for each of four faces.



As the towfish was lowered over the side of a boat, about 10 m away from the iceberg, the sonograph was displayed on a graphic recorder, and manually marked with the length of cable played out. A strong echo was received from the side of the iceberg, and the point at which this echo disappeared on the sonograph record was interpreted as indicating the depth at which the towfish had dropped beneath the lowest point of the iceberg. Interpreted iceberg profiles were made at four faces of the iceberg (Fig. 2). The overall draft of the iceberg was determined to be 18.0 ± 1.5 m.

Although the subsurface form of an iceberg can be obtained using side-looking sonar, this approach has some serious limitations. The technique is operationally very slow, and requires expensive logistic support, including a vessel. It also requires corrections for the tilt angle of the transducer from vertical and for the incident beam reaching grazing angle at the bottom of the iceberg, both of which are often difficult to estimate.

An iceberg's underwater shape clearly affects its stability, with most of an iceberg's ablation occurring beneath the waterline. As the iceberg melts it can become metastable so that only a slight perturbation is needed to initiate rolling. The wake from a small boat caused the iceberg under study to roll through 90° (Fig. 3). The underwater shape and the rolling process are important unknowns in understanding iceberg deterioration.

Although the measurements reported here are for a single, small iceberg, airborne radar shows considerable promise as a quick, safe, and accurate means of sounding icebergs. Interpretation of the radar return may be complicated by reflections from many facets within the iceberg; hence, the technique might be easier to apply to larger icebergs than to smaller ones. It might be most useful in relating the above-water form of icebergs to their subsurface draft on a statistical basis.

We thank T. Ridings, J. Johnstone, W. Winsor and J. Skidmore for assistance; A. Allan for encouragement. We also thank Memorial University Faculty of Engineering for the use of the side-scan sonar system.

JAMES R. ROSSITER
K. ANDREW GUSTAJTIS

Centre for Cold Ocean Resources Engineering,
Memorial University of Newfoundland,
St John's, Newfoundland,
Canada

Received 1 August; accepted 28 October 1977.

1. Kovacs, A. *Proc. 4th int. Conf. Port and Ocean Engineering under Arctic Conditions, St John's* (1977); *Proc. 1st int. Conf. Iceberg Utilization, Ames* (1977).
2. Swinbank, C. *Proc. 1st int. Conf. Iceberg Utilization, Ames* (1977).
3. Robin, G. de Q., Evans, S., & Bailey, J. T. *Phil. Trans. R. Soc., Lond. A* **265**, 437–505 (1969).
4. Goodman, R. H., Outcalt, E., Narod, B. B., & Clarke, G. K. C. *Proc. 4th Canad. Symp. Remote Sensing, Quebec* (1977).
5. Rossiter, J. R., & Gustajtis, K. A. *Proc. 1st int. Conf. Iceberg Utilization, Ames* (1977).
6. Smith, B. M. E. *Br. Antarctic Survey Sci. Rep.* **72** (1972).
7. Watts, R. D., & England, A. W. *J. Glaciol.* **17**, 39–48 (1976).
8. Benedict, C. P. *Proc. 2nd int. Conf. Port and Ocean Engineering Under Arctic Conditions, Reykjavik* 601–607 (1974); *ASME 76-Pet-103, Mexico City* (1976).

of an approaching continent meet an Andean-type margin of another continent, further lithospheric consumption is not possible due to the buoyancy of continental crust^{1,7} and either the crust of the underthrust continent is thickened, or two or more small plates (such as the Turkish, Aegean and Iranian plates) are formed and are pushed aside to permit further convergence, or both processes occur¹.

Molnar and Tapponnier^{2–4} have presented a similar analogy to explain certain tectonic features of Asia. They believe that the collision of India with Asia is responsible for the tectonic pattern we now see in Asia. Some crustal shortening is attributed to overthrusting in the Himalayas and in Tibet but the greatest amount of shortening is attributed to lateral movement of crustal blocks out of the path of the advancing subcontinent. In their model, India is analogous to a rigid indenter, Asia to a rigid-plastic solid, and large-scale strike-slip faults in Asia correspond to slip lines formed in response to plane indentation. They appealed to differently shaped rigid dies (Fig. 1a, b) in various parts of the collision zone to explain different features of the tectonic pattern of Asia. Of particular interest is the analogy they drew between the boundary of India in the western Himalayas and in Pakistan (Fig. 2a) and a wedge-shaped indenter (Fig. 1b).

A similar analogue model can be used to predict the behaviour of the Slave craton and the type of deformation that the Churchill craton underwent during collision. Here the Slave corresponds to a wedge-shaped rigid indenter, and the underthrust Churchill to a rigid-plastic medium (Figs 1b, 2b). The model predicts at least six major deformational features of the collision zone. These are now examined in turn to determine the success of the model.

According to the model (Fig. 1b) the northern and southern common boundaries of the Slave wedge and Churchill craton should be sinistral and dextral strike-slip faults respectively. The Bathurst and McDonald faults (Fig. 2b), which form the boundaries of the Slave wedge are major sinistral and dextral strike-slip fault systems that extend for hundreds of kilometres. They can be compared with the Quetta–Chaman and Karakoram faults in India (Fig. 2a). The model further predicts that the Slave wedge should behave as a rigid body relative to the Churchill. Hoffman *et al.*⁸ consider that relatively undeformed Archean rocks of the Slave province extend as far south as a wide belt of intensely lineated mylonites that occurs south of the McDonald fault in the Churchill province (Fig. 2b). Because the Slave granites are relatively undeformed and because the mylonites are apparently cut off by the McDonald fault, Hoffman *et al.* regard the mylonites as older than the fault and not genetically related to it. The present model, however, provides an alternative explanation for these relationships in terms of the different responses of the cratons to indentation; the rigid wedge (Slave) suffers relatively little deformation whereas the rigid-plastic medium (Churchill) is deformed intensely by widespread shearing and cataclasis over the areas indicated by the slip lines in Fig. 1b.

As indentation proceeds, crustal shortening is accomplished by local crustal thickening near the apex of the wedge, by sideways motion of crustal blocks (microplates) along strike-slip faults, or by both processes. The boundary between the Slave and Churchill provinces in the apical region is the Thelon metamorphic front⁹. Belts of cataclastic gneiss and mylonite occur in a wide zone roughly parallel to the front within the Churchill province¹⁰ and much of the eastern margin of the Slave is an extensive zone of shearing¹¹. The front is bounded on the Slave side by a regional negative gravity anomaly that is truncated by the Bathurst and McDonald faults and on the Churchill side by a coextensive positive anomaly (Fig. 3a). These paired anomalies have been interpreted as an edge-effect between juxtaposed crustal blocks of different mean density and thickness⁶; the Slave block is regarded as 'normal' crust whereas the Churchill crust is locally thickened and has a somewhat higher than normal density (Fig. 3b). Gibb and Thomas⁶ have interpreted the density discontinuity separating the blocks as a cryptic suture between collided cratons. The apical region of the Slave wedge is thus a wide zone of cataclasis where underthrusting of the Churchill has been inhibited by

Slave–Churchill collision tectonics

THE analogue model^{1–4} of plane indentation⁵ is applied qualitatively here to the proposed collision⁶ of the Slave craton (last major deformation ~ 2,500 Myr ago) and the western Churchill craton (last major deformation ~ 1,740 Myr ago) of the Canadian Shield. The model successfully predicts several large-scale deformational features of the collision zone. McKenzie¹ has drawn an analogy between the collision of the African plate (Arabian shield) with the Eurasian plate and the impaction of a hard die on a soft piece of metal. He concluded that when salients

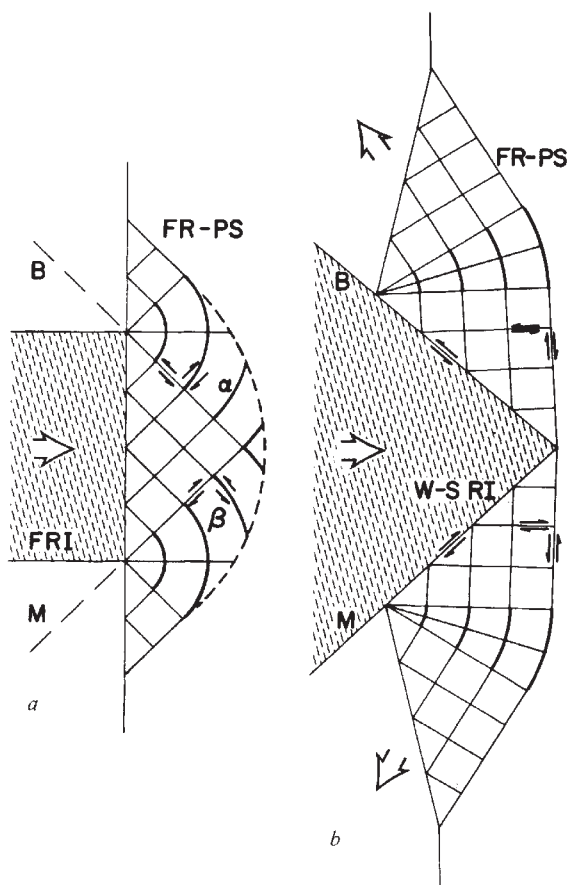


Fig. 1 Pattern of slip lines formed in the indentation of a semi-infinite flat rigid-plastic solid (FR-PS) by (a) a flat rigid indenter (FRI) and (b) a wedge-shaped rigid indenter (W-SRI) (modified from refs 1, 3, 5). Small arrows show sense of shear along α and β slip lines. Large arrows indicate direction of motion of indenters and displaced material of rigid-plastic solid. The analogous relationship of the Bathurst (B) and McDonald (M) faults to the models is also indicated.

buoyancy of continental crust resulting in localised crustal thickening as predicted by the model. It can be compared with the Pamir thrust belt in India (Fig. 2a).

Within the Churchill province strike-slip faults bounding microplates that may have developed north and south of the Slave wedge should display dextral and sinistral motion respectively (Figs 1b, 2b). Although the geology of these critical areas is relatively poorly known, a well-developed, unnamed, major fault mapped as a wide mylonite zone¹² (Fig. 2b) along with the Bathurst fault may define the boundaries of a northern microplate pushed sideways from the path of the advancing Slave wedge until such time as suturing was completed. A linear, negative magnetic anomaly coincides with the mylonite zone and extends north towards the coast and south to the Bathurst fault system with a total length of almost 300 km, confirming that the extent of the fault is much greater than shown on maps (Fig. 2b). The model predicts that the mylonite zone should be a dextral transcurrent fault (Fig. 1b) with a role similar to that of the Herat fault in Asia (Fig. 2a). The extension of the McDonald fault to the north-east (Fig. 2b) may be compared with the Talasso-Fergana fault in Asia (Fig. 2a) which has the same position relative to the wedge and the same sense of shearing. The proposed microplate south of the McDonald fault may be delimited to the east by either the sinistral Black Lake or Allan faults (Fig. 2b) and their northerly extensions or perhaps by an as yet unidentified fault corresponding to the Altyn Tagh fault in Asia (Fig. 2a). The Black Bay fault is another possible boundary fault; its geometry is consistent with the slip line pattern but its sense of movement is dextral.

North of the Bathurst fault reconnaissance mapping¹² has revealed several mylonite zones but not in sufficient detail to determine the strike-slip fault pattern predicted by slip line field theory (Fig. 1b). In the apical region, however, the effects of indentation are graphically illustrated by faults which dissect the Thelon magnetic high¹³ (Fig. 3c). Here major splays of the Bathurst¹³ and McDonald¹⁴ fault systems correspond to the pattern of slip lines formed by indentation of a flat die (Fig. 1a). Splay faults associated with the Bathurst fault system cut the magnetic high sinistrally¹³ as predicted by theory (β slip lines of Fig. 1a). A second set of faults almost normal to the first corresponds to α slip lines with a dextral sense of movement (Fig. 1a). South of the McDonald fault, the whole area as far south as Lake Athabasca is characterised by complex fault patterns¹⁵ which include wide mylonite zones, anastomosing faults related to the McDonald fault¹⁶ and to other major transcurrent faults, normal faults like those bordering the Nonacho basin¹⁷ and regional thrusts.

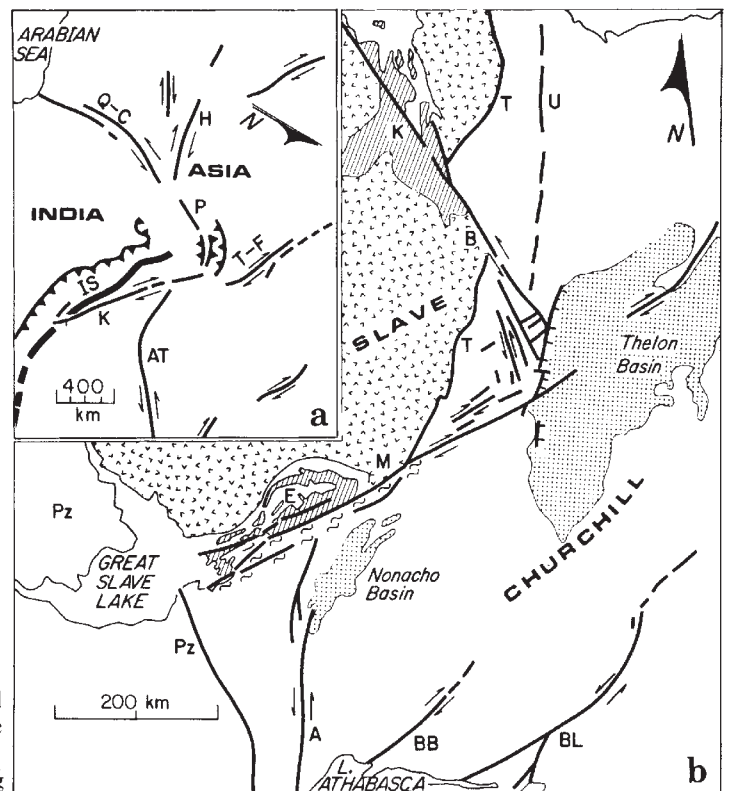


Fig. 2 Comparison of the major tectonic elements of (a) India and Asia in the vicinity of the Pamirs (P) (after ref. 3) with those of (b) the Slave-Churchill boundary zone in Canada (modified in part from refs 12-18). a, Q-C, Quetta-Chaman fault; H, Herat fault; K, Karakoram fault; T-F, Talasso-Fergana fault; AT, Altyn Tagh fault; IS, Indus suture. b, B, Bathurst fault; M, McDonald fault; T, Thelon front; U, unnamed mylonite zone and its extensions to north and south; A, Allan fault; BB, Black Bay fault; BL, Black Lake fault; K, Kilohigok basin; E, East Arm foldbelt. Ornament, v pattern, Slave province; waves, mylonite zone; diagonals, Aphebian basins; dots, Helikian basins; Pz, Paleozoic cover.

According to the wedge indentation model, rocks of the ancient Churchill continental margin originally at the apical contact region should be squeezed out and transported along the bounding faults of the indenter. Exotic breccias that have proved difficult to explain occur in the East Arm foldbelt in several fault-controlled zones associated with the McDonald fault¹⁸. Breccia fragments include both Aphebian sediments and Archean granite and range in size from large blocks tens of metres in length, to 'normal' breccia containing fragments a few centimetres in length set in a finer matrix. A possible explanation for these exotic breccias and other allochthonous fault-bounded blocks in the foldbelt is that they are chaotic fragments of the ancient Churchill

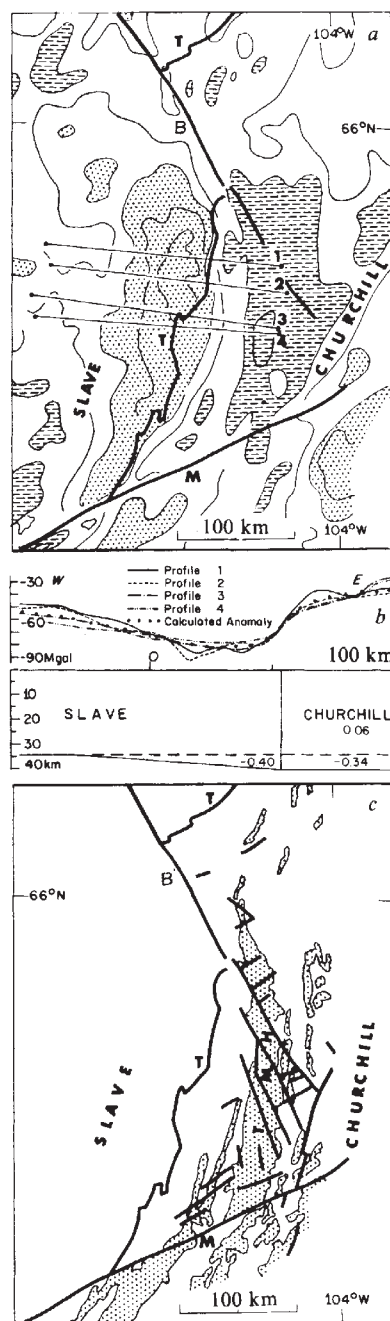


Fig. 3 a, Bouguer anomaly map of Slave-Churchill boundary zone (modified from ref. 6). Dots, anomalies < -60 Mgals; dashes, anomalies > -40 Mgals; 1-4, locations of profiles shown in b; B, M, T as in Fig. 2b. b, Profiles 1-4 showing Bouguer anomalies, calculated anomaly and interpreted crustal model (modified from ref. 6). Density contrasts in g cm^{-3} . c, Schematic magnetic map illustrating offsets of the Thelon magnetic high (TMH) by sinistral and dextral faults in the Slave-Churchill boundary zone (modified from ref. 13). Dots, total field magnetic anomalies $> 61,000$ nT; B, M, T as in Fig. 2b.

margin and craton that have been transported to their present locations in the Slave-Churchill boundary fault zone during the process of collision (indentation). Similar unexplained exotic breccias occur within the Aphebian Goulbourn Group rocks of the Kilohigok basin in the Bathurst fault zone¹⁹ and may have the same origin. Nappes have been described recently by Hoffman *et al.*⁸ in the East Arm foldbelt. They are up to 70 km long and 10 km wide⁸ and may have been ruptured and deposited during the same process; alternatively they may have been deposited by obduction of basin fill from the south during closure of the diminishing sea between the Slave wedge and Churchill craton. The presence of

compressional structures in the East Arm foldbelt and the juxtaposition, on either side of the foldbelt, of crustal blocks of distinctly different metamorphic history, metamorphic age, and deformational history tend to support a collision model rather than the aulacogen (failed rift) model proposed by Hoffman²⁰ in which tensional features should predominate and bordering cratons are likely to be similar.

These preliminary results suggest that the wedge indentation model or other variants of the indentation model may be useful for predicting deformational features of other proposed collision zones within the Canadian Shield²¹ as originally suggested by Molnar and Tappanier².

R. A. GIBB

Gravity and Geodynamics Division,
Earth Physics Branch,
Department of Energy, Mines and Resources,
Ottawa, Canada

Received 5 October; accepted 3 November 1977

- McKenzie, D. *Geophys. J. R. astr. Soc.* **30**, 109-185 (1972).
- Molnar, P. & Tappanier, P. *Science* **189**, 419-426 (1975).
- Tappanier, P. & Molnar, P. *Nature* **264**, 319-324 (1976).
- Molnar, P. & Tappanier, P. *Geology* **5**, 212-216 (1977).
- Hill, R. *The Mathematical Theory of Plasticity* (Oxford University Press, Oxford, 1950).
- Gibb, R. A. & Thomas, M. D. *Tectonophysics* **38**, 211-222 (1977).
- McKenzie, D. P. *Geophys. J. R. astr. Soc.* **18**, 1-32 (1969).
- Hoffman, P. F., Bell, I. R., Hildebrand, R. S. & Thorstad, L. *Pap. geol. Surv. Can.* **77** 1A, 117-129 (1977).
- Wright, G. M. *Mem. geol. Surv. Can.* **350** (1967).
- Davidson, A. *Spec. Pap. geol. Ass. Can.* **11**, 381-433 (1972).
- McGlynn, J. C. *Econ. Geol. Rep. geol. Surv. Can.* **1**, 71-76 (1970).
- Fraser, J. A. *Pap. geol. Surv. Can.* **63**-40 (1964).
- Thomas, M. D., Gibb, R. A. & Quince, J. R. *Can. J. Earth Sci.* **13**, 1244-1250 (1976).
- Fraser, J. A. *Pap. geol. Surv. Can.* **71**-38 (1972).
- Byers, A. R. *Spec. Publ. R. Soc. Can.* **4**, 40-59 (1962).
- Reinhardt, E. W. *Pap. geol. Surv. Can.* **69**-21 (1969).
- Burwash, R. A. & Baadsgaard, H. *Spec. Publ. R. Soc. Can.* **4**, 22-29 (1962).
- Reinhardt, E. W. *Pap. geol. Surv. Can.* **72**-25 (1972).
- Campbell, F. H. A. & Cecile, M. P. *Pap. geol. Surv. Can.* **76**-1A, 369-377 (1976).
- Hoffman, P. *Phil. Trans. R. Soc. Lond.* **A273**, 547-581 (1973).
- Gibb, R. A. & Thomas, M. D. *Nature* **262**, 199-200 (1976).

Inbreeding and dispersal in the great tit

IN populations which normally outbreed, matings between close relatives can result in a decrease in the viability and fertility of their offspring. Such inbreeding depression has been shown in a number of laboratory studies of insects¹, birds² and mammals³. We present here the first detailed evidence of inbreeding depression in a natural population and support for the hypothesis⁴ that one function of dispersal between birth and breeding sites is to reduce an individual's chance of inbreeding.

The paucity of long term studies on known individuals has meant that the frequency and consequences of inbreeding in natural populations are virtually unknown. Inbreeding has been reported in the great tit (*Parus major*)⁵⁻⁶, the yellow-eyed penguin (*Megadyptes antipodes*)⁷, the song sparrow (*Melospiza melodia*)⁸, the deermouse (*Peromyscus maniculatus*)⁹, in addition to probable cases in the gorilla (*Gorilla gorilla*)¹⁰ and olive baboon (*Papio anubis*)¹¹. In all cases, except the great tit⁶, no estimate was made of the relative frequency of inbreeding and the extent of inbreeding depression. Bulmer⁶ made an initial investigation of inbreeding using the great tit data collected from 1964-70 in Wytham Wood near Oxford and we have made a more extensive analysis, using the data collected in the period 1964-75.

From 1964 to 1975 there were 885 pairings in which the identities of both the male and the female were known. Pairs are formed both from birds born in the wood (residents) and from those born outside (immigrants). Details of the study area can be found elsewhere^{12,13}. The 885 pairs consisted of 194 matings between resident males and resident females, 239 between resident males and immigrant females, 158 between immigrant males and resident females and 294 between immigrant males and immigrant females. Of the total number, 13 (1.5%) were inbreeding pairs: five mother-son, seven

brother-sister (each pair from siblings in the same brood) and one aunt-nephew. If we include the same pairs breeding in 2 or more years, the total number of matings was 1,000, of which 16 (1.6%) were inbred. Two mother-son and one brother-sister pair each bred twice. The coefficient of inbreeding is

$$F = [(5 \times 0.25) + (7 \times 0.25) + (1 \times 0.125)] / 885 = 0.0035$$

It is probable that all cases of inbreeding between close relatives have been found⁶. The incompleteness of some of the pedigrees may have meant that a small, but undetected, number of inbreedings may also have occurred between more distant relatives, but such cases would produce only a slight increase in the coefficient of inbreeding.

From figures presented by Richdale⁷, we have also calculated the coefficient of inbreeding in the yellow-eyed penguin. In 244 pairings which produced 490 matings he found three brother-sister, one halfbrother-halfsister and three second cousin cases of inbreeding. The coefficient of inbreeding in this instance is

$$F = [(3 \times 0.25) + (1 \times 0.125) + (3 \times 0.0156)] / 244 = 0.0037$$

The frequency of close inbreeding (coefficient of kinship ≥ 0.125 (ref. 14)) is 1.6%. The figures are close to those found in the great tit: both species exhibit low levels of dispersal.

In laboratory studies the timing of reproduction, litter size, infant or fledging mortality^{2,3} and offspring viability and fertility^{3,15} may all be affected adversely by inbreeding. The laying date, clutch size, fledging success and number of offspring subsequently recorded breeding in the wood for all matings between relatives in the great tit are shown in Table 1. We have also calculated the dispersal distance from natal to breeding sites by all males, and those females mated to their brothers. In the cases of mothers and aunt mating with sons and nephew respectively, the appropriate measure of dispersal was the distance from previous outbreeding site to the subsequent inbreeding locality. For all these measures we have calculated the expected values from comparable outbred matings in which both parents were known (all second broods and relays have been excluded). Since there are yearly and age variations in laying date, clutch size¹² and dispersal (our work in preparation),

expected values are for each age-year category in question. We have only distinguished between first year and adult age classes, since sample sizes would otherwise have been small. The expected values for nesting success and the number subsequently recorded breeding in the wood have been calculated from the whole period 1964-75. We have excluded from the analysis all cases, in both inbreeding and outbreeding pairs, where breeding success was affected by predation, disturbance or experimental manipulation of brood size. For example, a great spotted woodpecker (*Dendrocopos major*) drilled a hole in the side of a nest box used by a mother-son pair in 1969. The incident occurred at hatching time and no offspring subsequently fledged; this pair has not been included in the calculation of nestling mortality.

On the three occasions when related birds bred in more than one year we have assumed that variables, such as laying date, are not independent measures. Consequently, in the analysis, we have averaged the observed and expected values for the 2 years.

The coefficient of kinship between mother and son is 0.25, as is, on average, that between full sibs, whereas that between aunt and nephew is 0.125. In our analysis we have included the latter example even though inbreeding depression would be expected to be less severe.

There is no evidence that inbreeding pairs produce smaller clutches than outbreeders (Wilcoxon test, $T = 32$, $n = 13$). Unexpectedly, however, the females mated with close relatives commenced laying before other females ($T = 15$, $n = 13$, $P < 0.05$). The reason for earlier laying is unknown, but may be connected with dispersal distance (our work in preparation). The result differs substantially from the laboratory study of *Peromyscus maniculatus*, in which siblings induced to mate, showed a delayed onset of reproduction when compared with unrelated pairs³.

Nestling mortality is higher among inbreeding than among outbreeding pairs ($T = 23$, $n = 15$, $P < 0.05$). We cannot distinguish between eggs which failed to hatch and offspring which died after hatching but before fledging. We refer collectively to the mortality as nestling mortality, which was, on average, 27.7% among inbreeders, whereas the expected mortality from equivalent sized clutches from outbreeders would be 16.2%.

Of the 86 inbred offspring which successfully fledged from mother-son and full sibling matings, five (5.8%) were subse-

Table 1 Observed and expected values for inbreeding pairs of great tits, 1964-75

Year	♂ Distance dispersed (m)		♀ Distance dispersed (m)		Day of first egg		No. of eggs		No. of fledglings		No. recovered breeding	
	O	E	O	E	O	E	O	E	O	E	O	E
Mother-son												
{ 1965	14	466.7	0	77.7	122	126.9	8	8.38	8	6.72	0	0.84
{ 1967	—	—	—	—	129	121.5	7	8.10	4	5.97	0	0.36
{ 1968	120	382.7	50	90.5	115	120.1	6	8.35	3	4.83	1	0.18
{ 1969*	—	—	—	—	133	123.2	7	8.28	(0)	—	—	—
1971	91	623.9	110	91.7	124	126.9	8	8.41	3	6.72	0	0.18
1973	110	453.1	105	55.9	121	123.8	9	8.48	8	7.53	0	0.84
1974	0	378.5	0	64.0	132	131.8	10	7.06	6	8.43	0	0.73
Brother-sister												
1966	151	690.3	151	826.4	127	133.1	10	8.29	6	8.43	0	0.73
{ 1967	500	340.1	500	912.5	118	121.5	10	8.10	6	8.43	0	0.73
{ 1968	—	—	—	—	116	120.1	10	8.35	9	8.43	0	0.79
1970	237	463.1	237	944.7	123	126.8	10	9.83	9	8.43	0	0.79
1971	970	623.9	970	749.7	131	126.9	7	8.41	4	5.97	1	0.36
1972	1,486	453.1	1,486	758.2	126	133.1	9	8.29	6	7.53	1	0.73
1972	632	623.9	632	749.7	119	122.4	8	8.03	7	6.72	2	0.65
1972	356	623.9	356	749.7	116	122.4	9	8.03	7	7.53	0	0.65
1977	835	—	835	—	—	—	6	—	6	4.83	—	—
1977	380	—	380	—	—	—	8	—	6	6.72	—	—
Aunt-nephew												
1966	41	340.1	36	90.6	128	133.1	9	8.29	8	7.53	1	0.84

Observed (O) and expected values (E) for inbreeding pairs 1964-75. Pairs breeding in more than 1 yr are bracketed together. Two brother-sister pairs recorded in 1977 were used only in the analysis of nestling mortality.

* Breeding disturbed by woodpecker at time of hatching.

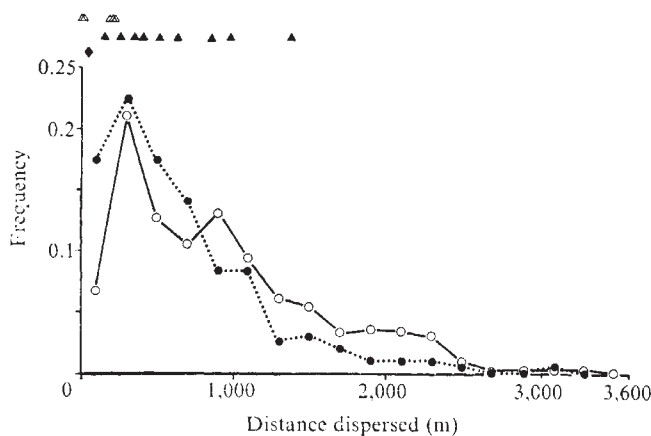


Fig. 1 Distance dispersed by outbreeding individuals between natal site and site of first breeding. The distance localities of inbreeding are marked for males only, when mated to mothers and aunt, and for both males and females, moving the same distance, in brother-sister pairs. Dispersal:; males; —, females. Inbreeding: Δ , mother-son; $\blacktriangle$, brother-sister; $\blacklozenge$, aunt-nephew.

quently recovered breeding in the wood. If we include the one produced by the aunt-nephew pair, six (6.4%) returned from 94 offspring. The expected number from normal broods were respectively 8.56 (10.0%) and 9.40 (10.0%). If we assume that the number of recoveries fits a Poisson distribution, then the probabilities of obtaining five or less from 86 and six or less from 94 fledglings are 0.145 and 0.174 respectively. Mortality after fledging is not significantly greater than expected.

The six inbred offspring (three males and three females) subsequently outbred and produced an average clutch of eight eggs. The nestling mortality of their offspring was 16.7%. None of the offspring of mother-son or brother-sister pairs survived longer than one breeding season, whereas that of the aunt-nephew pair survived 2 years.

The distance dispersed between natal and breeding site may affect the chances of mating with a close relative. The patterns of dispersal of normal juvenile males and females, together with the locations of inbreeding are shown in Fig. 1. Females disperse further than males in their first year⁶; but subsequently, there is little movement and individuals tend to retain the same territory from one year to the next (our work in preparation). These patterns of dispersal lead to the expectation that sons mating with their mothers would have dispersed substantially less than expected. In sibling matings, brothers should move more than the expected median distance or sisters less than expected, or both. These predictions are supported by the data, although formal testing is inappropriate because of the lack of independence of dispersal between related inbreeding individuals. On average, sons mated with their mothers moved only 67 m from natal to breeding site, 85% less than the expected distance (461 m). Males mated to their sisters dispersed 13.5% more than expected, 619 m as opposed to 545 m, whereas sisters moved less than expected (24%), 619 m rather than 812 m.

From the distribution of dispersals it is possible to calculate the expected number of inbreedings from the model produced by Bulmer⁶. For close relatives these are: mother-son, 6.2 (observed 5); brother-sister, 7.35 (observed 7); and father-daughter, 2.2 (observed 0). More examples are necessary, before it is possible to estimate whether fewer inbreedings are occurring than would be expected.

We have shown that inbreeding pairs have a lower breeding success, resulting from higher nesting mortality, than normal pairs and that the chances of inbreeding vary with dispersal distance. Since the avoidance of such matings is one factor influencing dispersal, it is possible that the asymmetry in movement between males and females in many sedentary species

is maintained by the detrimental effects of inbreeding^{4,16-18}.

We are grateful to Tony Martin and Peter Croydon for computer programming support. We thank Michael Bulmer, Tim Clutton-Brock and members of the Population Biology Group at Sussex for comments, and John Maynard Smith for much help and encouragement. This work was supported by the SRC.

PAUL J. GREENWOOD
PAUL H. HARVEY

School of Biological Sciences,
University of Sussex,
Falmer, Brighton, Sussex, UK

CHRISTOPHER M. PERRINS

Edward Grey Institute,
Department of Zoology,
South Parks Road, Oxford, UK

Received 13 October; accepted 9 November 1977.

1. Maynard Smith, J. *J. Genet.* **54**, 261-279 (1956).
2. Sittman, K., Abplanalp, B. & Fraser, R. A. *Genetics* **54**, 371-379 (1966).
3. Hill, J. L. *Science* **186**, 1042-1044 (1974).
4. Greenwood, P. J. & Harvey, P. H. *J. Anim. Ecol.* **45**, 887-898 (1976).
5. Kluijver, H. N. *Ardea* **39**, 1-135 (1951).
6. Bulmer, M. G. *Heredity* **30**, 313-325 (1973).
7. Richdale, L. E. *A Population Study of Penguins* (Clarendon, Oxford, 1957).
8. Nice, M. M. *Trans. linn. Soc. N.Y.* **4**, 1-247 (1937).
9. Howard, W. E. *Contr. Lab. vert. Biol. Univ. Mich.* **43**, 1-50 (1949).
10. Harcourt, A. H., Stewart, K. S. & Fossey, D. *Nature* **263**, 226-227 (1976).
11. Packer, C. thesis, Univ. Sussex (1977).
12. Lack, D. *Population Studies of Birds* (Clarendon, Oxford, 1966).
13. Perrins, C. M. *J. Anim. Ecol.* **34**, 601-647 (1965).
14. Jacquard, A. *Genetic Structure of Populations* (Springer, Berlin, 1974).
15. Dobzhansky, T. *Genetics of the Evolutionary Process* (Columbia University Press, New York, 1970).
16. Verner, J. *Bird-Banding* **42**, 92-98 (1971).
17. Clutton-Brock, T. H. & Harvey, P. H. in *Growing Points in Ethology* (ed. P. P. G. Bateson & R. A. Hinde), 195-237 (Cambridge University Press, London, 1976).
18. Greenwood, P. J. & Harvey, P. H. *Ibis* **119**, 528-531 (1977).

Cortical magnification factor predicts the photopic contrast sensitivity of peripheral vision

THE area of visual cortex devoted to the analysis of a constant-size region in the visual field diminishes progressively for more peripheral locations. The change is described quantitatively by the cortical magnification factor, which indicates the linear extent of cortex in mm corresponding to one degree of visual angle at various eccentricities (angular distances from the middle fovea). The human cortical magnification factor has been estimated by Cowey and Rolls¹ from the data of Brindley and Lewin², who mapped the phosphenes (sensations of light) caused in the lower nasal visual field by electrical stimulation of the human visual cortex. Building on these results, we have studied the spatial contrast sensitivity functions in man at various eccentricities. We used two methods: in one the retinal image sizes of the test gratings were kept similar at different eccentricities and in the other, the calculated sizes of cortical projections of the test gratings were made similar at different locations. Our results indicate that peripheral contrast sensitivity and acuity are inferior to foveal performance, because of the diminished cortical projection area.

Four experienced subjects fixated a small point of light binocularly in a dark room; a bite board aided steady fixation. The contrast sensitivity at various eccentricities along the nasal horizontal and inferior vertical visual-field meridians of the right eye was measured by determining the inverse of the threshold contrast for each of a series of sinusoidal gratings. The gratings were generated on a white cathode-ray screen³, masked to expose only half of a round field 16 cm in original diameter (see inset, Fig. 1a). The screen was perpendicular to the visual axis if viewed directly and at the same distance as the fixation point; the screen was not visible to the left eye.

The average luminance of the screen was continually

11 cd m^{-2} and the diameter of the natural pupil varied from 6 to 7 mm. (No artificial pupil was used because it limits the visual field.) The photopic average retinal illuminance was thus about 365 Trol (1 troland = $1 \text{ cd m}^{-2} \times \text{pupillary area in mm}^2$). On our display this corresponds to about 1,030 scotopic Trol, indicating a photopic adaptation level.

Thresholds were determined by using a computer-controlled forced-choice method: the subject received two sound signals and had to decide during which signal a grating was presented. Another sound indicated the wrong choices. After each block of four consecutive correct responses, contrast was decreased by 2 dB and every wrong choice led to a contrast increment of 2 dB: the contrast required for a probability of 0.84 of correct response was estimated from the turning points as described by Wetherill and Levitt⁴. The critical stimulus to be detected was a stationary vertical grating with a non-zero contrast for 0.5 s; the control stimulus was otherwise identical but had a zero contrast. To avoid the bias caused by the Troxler effect (fading of stimuli in peripheral vision) the subject initiated each threshold trial only when the display was clearly visible. Control experiments showed that results essentially similar to those of the detection method are obtained from a two-choice recognition of grating orientation or direction of movement.

Figure 1 shows the results for subject V.V. in the lower half of the vertical meridian; the results of other subjects and of the nasal horizontal visual field were similar. In the experiments of Fig. 1a the display was viewed at a constant distance of 458 cm. The retinal images of the gratings subtended $1 \times 2^\circ$. At increasing eccentricities, contrast sensitivity and visual acuity (assessed from the resolution of high-contrast gratings) decreased as similar decelerating functions of eccentricity.

Figure 1b presents results for the same eccentricities, but in this case both the spatial frequency and the size of the gratings were scaled by the inverse of the cortical magnification factor estimated from the points calculated by Cowey and Rolls⁵. Thus, the cortical topographical representations of the gratings became equal in size at various eccentricities. For the series of eccentricities: 0, 1.5, 4, 7.5, 14 and 30° the following magnification factors were assumed: 7.75,

5.25, 3.44, 2.31, 1.24 and 0.49 mm. The scaling of gratings by the inverses of these values was accomplished by keeping the gratings constant and moving the subject closer to the display; the effects of the changes of accommodation on acuity were partially compensated by corrective lenses. The viewing distances for the semicircular gratings with radius 8 cm were 458, 312, 204, 137, 73.3 and 28.3 cm. Thus the gratings subtended $1 \times 2^\circ$ at $E=0$, and $15.8 \times 31.6^\circ$ at $E=30^\circ$, with corresponding intermediate retinal sizes. Assuming that the magnification factors were correct, the crescent-shaped cortical projection fields had a constant size of about $7.75 \times 15.5 \text{ mm}$ and the fields stimulated at neighbouring eccentricities were tangential to each other.

After scaling, the widely separate contrast sensitivity functions of different eccentricities collapsed to an apparently single function, as in Fig. 1b. This is surprising, because different types of retinal ganglion cells may dominate in peripheral and foveal visual systems, and the projections reach several locations in the brain⁶. A good correlation between conventional measures of visual acuity and the magnification factor^{1,6} can be explained by postulating a homogeneous detection system subserving acuity, but there is evidence that different types of mechanisms mediate the detection of low and high retinal spatial frequencies⁷.

A minor deviation from the unity of the scaled functions occurred at high spatial frequencies (Fig. 1c), but this was expected for optical reasons. The optical limitations of contrast transfer with a pupil size of 6–7 mm are considerable at high retinal spatial frequencies, but they should not affect the transfer of low retinal spatial frequencies⁸ that represent high 'cortical' spatial frequencies in peripheral vision. Figure 1c shows that sensitivity to the highest 'cortical' spatial frequencies improves somewhat at greater eccentricities, as would be expected if optical factors limited performance at small eccentricities. The small overall decrease of sensitivity from $E=0^\circ$ to $E=1.5^\circ$ reflects the Troxler effect and also our uncertainty regarding the exact value of the foveal magnification factor.

We conclude that the foveal contrast sensitivity function can be generalised for any part of the visual field when two structurally determined adjustments are made: one is based

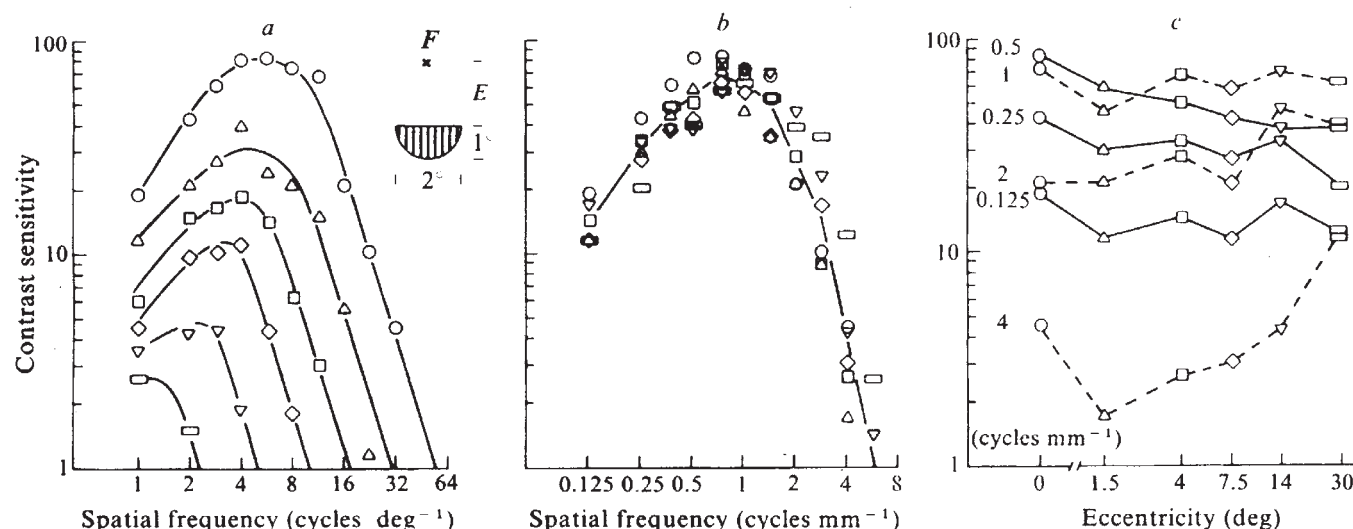


Fig. 1 Contrast sensitivity as a function of spatial frequency and eccentricity (E) in the lower visual field. *a*, The retinal images of various sinusoidal test gratings were similar at different eccentricities as shown in the inset. Spatial frequency indicates the bar density of gratings in cycles per degree of visual angle and eccentricity refers to the angular distance between the fixation point (F) and the nearest edge of the semicircular vertical grating. Curves represent values of E as follows: $\circ$, 0° ; $\triangle$, 1.5° ; $\square$, 4° ; $\diamond$, 7.5° ; ∇ , 14° ; $\blacksquare$, 30° . *b*, Cortical projections of the test gratings were similar at different locations. The symbols are as in *a* regarding eccentricity in the visual field. Spatial frequency indicates the estimated bar density in cycles per mm on the primary visual cortex. *c*, Part of the data of *b* replotted to reveal the expected improvement of sensitivity as a function of E at high cortical spatial frequencies.

on the cortical magnification factor and the other on the modulation transfer function of the eye optics. The result suggests that, using suitable temporal conditions, a picture can be made equally visible at any eccentricity by scaling its size by the magnification factor, because the contrast sensitivity function represents the spatial modulation transfer function of the visual system for near-threshold contrasts. The results by Anstis⁹ on the recognition of letters in peripheral vision agree with this generalisation.

The invariance of the photopic contrast sensitivity functions of different but equally large regions of the visual cortex indicates a striking functional homogeneity of image processing machinery at central levels of the visual system. This functional homogeneity is not normally evident because it is masked by various early transformations of the signals entering the cortex.

We thank the Finnish Academy and the Finnish National Fund for Research and Development for support to V.V., and Pekka K. Lehtiö and Pentti Laurinen for design of apparatus and programmes.

JYRKI ROVAMO

*Institute of Physiology,
University of Helsinki,
Siltavuorenpenger 20a, SF00170 Helsinki 17, Finland*

VEIJO VIRSU
RISTO NÄSÄNEN

*Department of General Psychology,
University of Helsinki,
Ritarikatu 5, SF00170 Helsinki 17, Finland*

Received 14 July; accepted 9 November 1977.

1. Cowey, A. & Rolls, E. T. *Exp Brain Res.* 21, 447-454 (1974).
2. Brindley, G. S. & Lewin, W. S. J. *Physiol., Lond.* 196, 479-493 (1968).
3. Virsu, V., Nyman, G. & Lehtiö, P. K. *Perception* 3, 323-336 (1974).
4. Wetherill, G. B. & Levitt, H. *Br. J. math. stat. Psychol.* 18, 1-10 (1965).
5. Lennie, P. *Nature* 266, 496 (1977).
6. Drasdo, N. *Nature* 266, 554-556 (1977).
7. Kulikowski, J. J. & Tolhurst, D. J. *J. Physiol., Lond.* 232, 149-162 (1973).
8. Green, D. G. *J. Physiol., Lond.* 207, 351-356 (1970).
9. Anstis, S. M. *Vision Res.* 14, 589-592 (1974).

Is *Acanthaster planci* an *r*-strategist?

THE debate over the cause of plagues of the crown-of-thorns starfish, *Acanthaster planci*, has been as long as the plagues themselves (for reviews see refs 1-3). Most recently, two lines of reasoning based on ecological concepts have been used to suggest that the causes of the population explosions should be sought in some abnormal event, and thus man has been implicated²⁻⁴. I believe that a more detailed consideration of the ecology of *A. planci* does not support this view; rather, it suggests violent population fluctuations without the assistance of man's activities.

It has been predicted² that complex ecosystems such as coral reefs should be resistant to perturbations because of the supposed stabilising effect of high trophic complexity⁵. It is now more widely thought⁶⁻⁸ that the stability often apparent is a consequence of a highly predictable environment, which, it is believed, on coral reefs⁷ allows a great diversity of organisms to specialise on the large number of spatial and biotic niches created there. Such specialised species are characterised by small, stable population sizes and low reproductive rates, and are long-lived.

Grassale's descriptive framework⁷ coincides with Pianka's visualisation of an *r*-*K* continuum⁹, the two extremes representing opposing bionomic strategies for survival in the face of natural selection: *r*-strategist species (or 'opportunists'^{7,10}) tend to exploit unpredictable and/or

short-lived habitats¹¹ by devoting energy and resources to maximising reproductive potential; *K*-selected species (or 'specialists'⁷) exhibit conservative strategies for survival in predictable environments, and are adapted against pressures from predators and inter-specific competition.

Examining some of the important sessile components of the coral reef ecosystem within this framework, it seems that most corals¹² and sponges¹³, and several other organisms¹⁴ exhibit conservative strategies in view of their massive size, long life, and adaptations for inter-specific competition, although some¹⁵ are adapted to exploiting unpredictable habitats.

Where does *A. planci* fit into this picture? Its bionomic strategy is indicated by features of its ecology and life history⁹ which are shown in Table 1. There are several other coral reef echinoderms whose population distributions, like that of *A. planci*, are patchy but locally abundant, and they have also been examined within this framework, with the limited data available.

A. planci has probably a greater range within the coral reef environment than any other echinoderm, being found virtually wherever there is coral. Most other echinoderms are restricted by their food requirements to narrow zones. Studies in the Red Sea³ have indicated some density-dependent predation by fish, but it seems that when population densities become very high, as on the Great Barrier Reef³, they escape the control of predators; migration then becomes important. Population sizes vary markedly in time and place¹⁻³; normal population densities are rather low, and dense populations are uncommon¹⁶. Dense populations can persist aided by the virtual absence of inter-specific competition for coral prey from other reef inhabitants: most fish or other coral predators only graze polyps sub-lethally or kill small parts of colonies^{17,18}. None, except perhaps donkey-fish (*Bulbometopon muricatus*) locally (my observations), would affect significantly the food available to *A. planci*: it is apparently exploiting what is virtually an ecological vacuum.

Whereas coral is killed by *A. planci* and takes some years to be replaced, the nature of the food resources of many other common coral reef echinoderms (for example, the starfish *Linckia multiflora* (my observations) and *L. laevigata*¹⁹—substrate feeders, primarily on coralline algae, or the holothuroid *Holothuria atra*—a sediment feeder²⁰) is such that it is continuously being renewed: in their more restricted zones, they compete intra- and inter-specifically for food, near the carrying capacity of their environment. Thus, although the potential for rapid increase in population numbers is present in many coral reef echinoderms, by virtue of the large numbers of larvae produced, adult numbers of most species would seem to be limited by restricted zonation and food availability.

By contrast, *A. planci* feeds selectively at low densities, and when its population density increases, it adapts to unavailability of preferred species²¹, feeding on all scleractinian corals^{1,2}. Also, being relatively mobile, it may exploit extensive areas of reef when its food supply becomes locally depleted², and in these ways it overcomes the problem of intra-specific competition for food.

Table 1 shows that many of its characteristics indicate *A. planci* to be strongly *r*-selected, while other common coral-reef echinoderms exhibit more moderate strategies. But some aspects of *A. planci*'s reproductive strategy, being different from those predicted for *r*-selected terrestrial animals, seem to contradict this view. Upon examination, it becomes clear that different considerations must be applied. We may predict that *r*-selection would favour the production of large numbers of larvae, and most of the commoner species of coral reef asteroids, including *A. planci*, produce many small, planktotrophic, pelagic larvae²². Their survival, however, is probably very

Table 1 Bionomic strategy of *Acanthaster planci* and other common coral reef echinoderms as indicated by life-history characteristics

Correlates of <i>r</i> - and <i>K</i> -selection*	Characteristics of <i>A. planci</i>	Strategy indicated <i>r</i> <i>K</i>	Other common coral reef echinoderms†	Strategy indicated <i>r</i> <i>K</i>
Variability/predictability of environment	Most zones on coral reefs, 1–50 m depth	+	Predictable, restricted zonation	+
Mortality	Density-dependent at low densities, but escapes this control at higher densities; migration important	+	Not known	— —
Survivorship	Extensive mortality of larvae	+	Extensive mortality of larvae	+
Population size:				
—variability in time	Variable	+	Less variable than <i>A. planci</i> ?	?
—relation to carrying capacity	Usually low	+	Often high	+
—saturation of community	Ecological vacuum	+	More saturated	+
—recolonisation	Migration of larvae important	?	Larvae migrate; asexual reproduction occurs in some species	? +
Competition:				
—inter-specific	Negligible	+	Varies with species	— —
—intra-specific	Low except at very high densities	+	Limits numbers locally	
Characteristics predicted as favoured by selection:				
—rate of development	Fast	+	Relatively slow	+
— r_{max}	Potentially very high	+	Potentially very high	+
—age of sexual maturity	1–3 yr		Not known, but probably comparable	+
—body size	Large compared to other coral reef echinoderms		Varies with species, but all possess anti-predator defences	+
—reproductive strategy	Repeated		Repeated	+
Longevity	8 yr		Not known, but probably comparable	+

Data sources available from author. +, Denotes clearly indicated *r* or *K* strategy; ?, strategy not clear.

*After Pianka⁸.

†Details of individual species are omitted for sake of brevity. Actual species compared are: *Culcita novaeguineae*, *Linckia multiflora*, *L. laevigata*¹⁸, *Diadema setosum* and *Holothuria atra*¹⁹.

uncertain^{23,24}, particularly in the case of *A. planci*, whose high survival rates of larvae may result from unusual levels of water salinity and temperature and food supply (ref. 25 and R. Pearson quoted in ref. 6). The uncertainty in recruitment of juveniles must be counteracted by a shift towards *K*-selection in the adult, reflected in increased longevity, as demonstrated in some marine fishes²⁷: in *A. planci* the life span is believed to be about eight years²⁸. Anti-predator defences (which complement increased longevity) are also present in most common coral reef echinoderms; *A. planci* has spines and a cryptic behaviour, as do many echinoids, and other asteroids have tough skins or heavily-calcified skeletons²². Large size may act in the same way¹¹.

It would be unrealistic to attempt to model mathematically the population dynamics resulting from the bionomic strategy of *A. planci* as outlined above, owing to its apparently 'fugitive' nature (local populations usually have a unimodal age distribution²⁴).

Clearly *A. planci* is not a 'classical' *r*-strategist in the sense of having an unstable or short-lived habitat, but its life-history characteristics indicate a coherent strategy for exploiting what is, surprisingly, a virtual ecological vacuum. In some ways its strategy resembles that of certain insect pests, for example, the spruce budworm²⁹, which conform broadly to *r*-strategies, but exploit longer-lived habitats.

In view of the very high reproductive potential (r_{max}) and the possibly long period before environmental conditions favour massive larval survival, the large fluctuations of local populations which characterise *r*-strategists¹¹ can only be exaggerated in the case of *A. planci*. Thus it is inappropriate to regard the low population levels usually observed^{2,3,16} as 'normal' and high population levels as 'abnormal': periodic plagues are inherent to this mode of living and may well be necessary for survival of the species. It is unfortunate for many fish and other reef inhabitants that such plagues result in the destruction of their habitat³⁰—the effects are comparable to those of destroying the canopy of a tropical rain forest. Coral reefs

are also valuable to man, and it is reflection of his subjective viewpoint as a *K*-strategist, striving to maintain his resources, that the terms 'plague' and 'population explosion' tend to be imbued with emotion.

This paper arose from work carried out during the crown-of-thorns research project funded by the Ministry of Overseas Development. I thank John Lawton and Rupert Ormond for stimulating discussions.

RICHARD J. MOORE

Cambridge Research Group, PO Box 24,
Port Sudan, Democratic Republic of Sudan

Received 21 March; accepted 18 October 1977.

1. Branham, J. M. *Bioscience* 23, 219–226 (1973).
2. Endean, R. in *Proc. Second Int. Coral Reef Symp.* 1, 563–576 (Great Barrier Reef Committee, Brisbane, 1974).
3. Ormond, R. F. G. & Campbell, A. C. in *Proc. Second Int. Coral Reef Symp.* 1, 595–619 (Great Barrier Reef Committee, Brisbane, 1974).
4. Randall, J. E. *Biotropica* 4, 132–144 (1972).
5. MacArthur, R. H. *Ecology* 36, 535–536 (1955).
6. May, R. M. *Math. Biosci.* 12, 59–79 (1971).
7. Grassle, J. E., in *Biology and Geology of Coral Reefs* 2, (eds Jones, O. A. & Endean, R.) 247–270 (New York, 1973).
8. Sanders, H. L. *Am. Nat.* 102, 243–282 (1968).
9. Pianka, E. R. *Am. Nat.* 104, 592–597 (1970).
10. MacArthur, R. H. *Am. Nat.* 94, 25–34 (1960).
11. Southwood, T. R. E., May, R. M., Hassel, M. P. & Conway, G. R. *Am. Nat.* 108, 791–804 (1974).
12. Connell, J. H. in *Biology and Geology of Coral Reefs* 2, (eds Jones, O. A. & Endean, R.) 205–245 (New York, 1973).
13. Jackson, J. B. C. & Buss, L. *Proc. natn. Acad. Sci. U.S.A.* 72, 5160–5163 (1975).
14. Burkholder, P. R. in *Biology and Geology of Coral Reefs* 2, (eds Jones, O. A. & Endean, R.) 117–182 (New York, 1973).
15. Loya, Y. *Nature* 259, 478–480 (1976).
16. Vine, P. J. *Nature* 228, 341–342 (1970).
17. Robertson, R. *Pacific Sci.* 24, 43–54 (1970).
18. Randall, J. E. in *Proc. Second Int. Coral Reef Symp.* 1, 159–166 (Great Barrier Reef Committee, Brisbane, 1974).
19. Laxton, R. *Biol. J. Linn. Soc.* 6 (1), 47–64 (1973).
20. Bakus, G. J. in *Biology and Geology of Coral Reefs* 2, (eds Jones, O. A. & Endean, R.) 325–367 (New York, 1973).
21. Ormond, R. F. G., Hanscomb, N. J. & Beach, D. H. *Mar. Behav. & Physiol.* 4, 93–105 (1976).
22. Yamaguchi, M. *Biotropica* 7 (1), 12–23 (1975).
23. Thorson, G. *Biol. Rev.* 25, 1–45 (1950).
24. Yamaguchi, M. *Micronesica* 9 (2), 207–212 (1973).
25. Lucas, J. S. *Micronesica* 9 (2), 197–203 (1973).
26. Pearson, R. Report on research sponsored by the Advisory Committee on research into the Crown-of-thorns starfish (Australian Government Publishing Service, Canberra, 1975).
27. Murphy, G. I. *Am. Nat.* 102, 391–403 (1968).
28. Cheshier, R. H. *Science* 165, 280 (1969).
29. Morris, R. F. (ed.) *Mem. Entomol. Soc. Can.* No. 31 (1963).
30. Endean, R. & Stabium, W. *Atoll. Res. Bull.* No. 168 (1973).

Requirement for cell differentiation in *Dictyostelium discoideum*

THE generation of cell diversity during the development of multicellular organisms involves a series of decisions between alternative pathways of differentiation^{1,2}. In many instances the fate of a cell depends on its position, so that pattern formation and cell differentiation can become facets of the same process³. These phenomena occur in a simple form in the slime mould *Dictyostelium discoideum* where there are only two major pathways of differentiation—one leading to stalk cells and the other to spores. On starvation the individual amoebae aggregate chemotactically to form multicellular masses. Somewhat later the aggregates can be shown to consist of two distinct cell types: prestalk cells which are localised in the anterior and prespore cells in the posterior of the migrating slug⁴. We are interested in how the formation of these cell types is controlled. One approach to this problem involves experimentally inducing differentiation of cells that are somehow prevented from completing normal development. Starving cells plated on agar with a Cellophane overlay aggregate but remain constrained as a monolayer and do not progress to stalk or spore. But addition of 1–5 mM cyclic AMP to the agar results in efficient induction of stalk cell formation provided that a high cell density is used^{5–7}. At low cell density stalk cells do not form unless a low-molecular weight

factor, released by cells at high density, is also provided⁷. We show here that although no mature spores form in conditions of high cell density giving a high yield of stalk cells large numbers of prespore cells do appear. Many of these later transform into stalk cells. Prespore cells do not arise at low density but the density requirement for prespore cell formation, unlike that of stalk cells, cannot be efficiently replaced by the low-molecular weight factor. Five mutants isolated by virtue of their ability to form mature spores under Cellophane likewise do not form spores at low cell density even in the presence of the 'helping' factor. Thus entry of developing *Dictyostelium* cells into the prespore pathway apparently requires some cell interaction different from or additional to that required for entry into the stalk pathway.

Our initial observations, made by electron microscopy, were concerned with the induction of prespore vacuoles (PSVs), which most workers consider to be specific markers of prespore cells^{8–12}. HMI cells were plated at high cell density on agar and overlaid with Cellophane as described in the legend to Fig. 1. In the absence of added cyclic AMP few PSVs and no stalk cells were observed (Fig. 1). When 5 mM cyclic AMP was included in the agar, large numbers of PSVs were induced. They first appeared by 15 h of induction, reached a peak number by 24 h and thereafter declined in numbers as mature stalk cells accumulated. The latter were never seen to contain a PSV.

The experiment described above was rather time consuming and so in the following experiments we identified prespore cells by their specific fluorescent staining with absorbed anti-spore serum^{13,14}. Figure 2b again demonstrates the dramatic induction of prespore cells by 5 mM cyclic AMP at high cell density. This result is in close qualitative agreement with that obtained by electron microscopy. Precedents for these results are to be found in the reports of the occurrence (but not the induction) of prespore cells in the presence of 1 mM cyclic AMP^{15,16}. The kinetics of prespore induction shown in Figs 1 and 2b suggest that

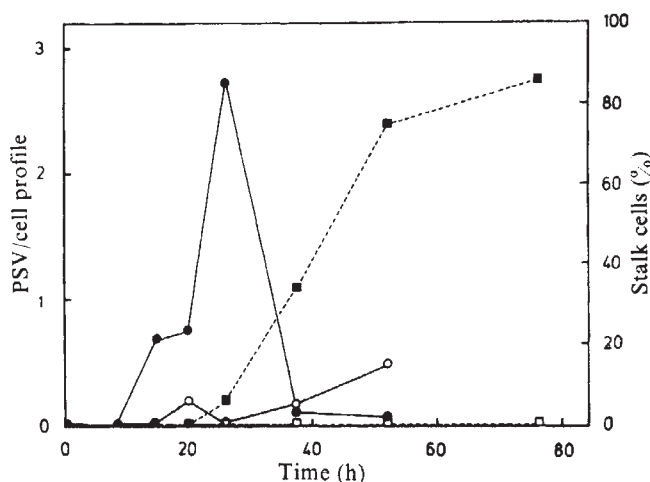
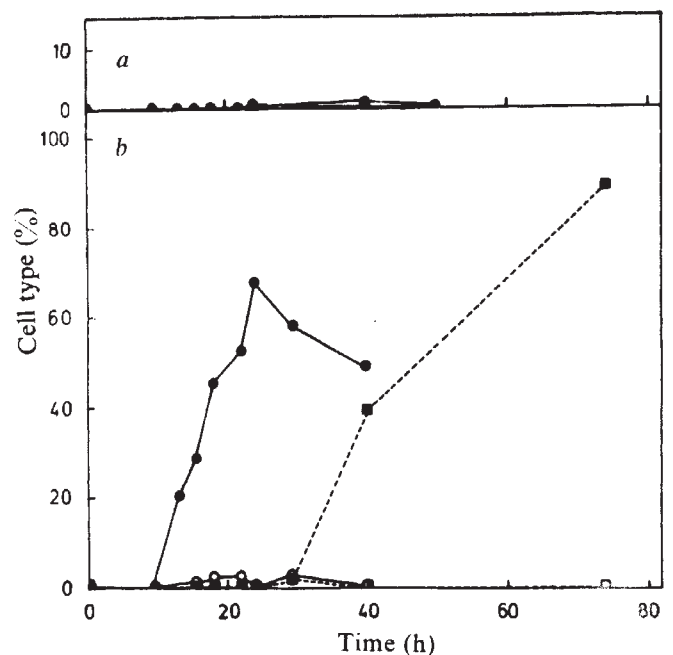


Fig. 1 Induction of prespore vacuoles and stalk cells by cyclic AMP in cells plated at high density under Cellophane. ○, Prespore vacuoles, no cyclic AMP; □, stalk cells, no cyclic AMP; ●, prespore vacuoles + cyclic AMP; ■, stalk cells + cyclic AMP. HMI cells (HMI is an acriflavin resistant strain derived from V12 M2) were grown on SM agar with *Klebsiella aerogenes* and collected in sterile KK_2 buffer (16.5 mM, KH_2PO_4 , 3.8 mM K_2HPO_4 , 2 mM MgSO_4 , pH ~ 6.2) with 60 mM KCl added. They were pelleted at 200g for 3 min, washed three times in KK_2 and finally once in 5% Bonner's salts (100% is $0.6 \text{ g l}^{-1} \text{ NaCl}$, $0.75 \text{ g l}^{-1} \text{ KCl}$, $0.3 \text{ g l}^{-1} \text{ CaCl}_2$). Cells at a density of $2.5 \times 10^5 \text{ cm}^{-2}$ were spread on 2 ml of 1.5% agar (Oxoid L28) containing 5% Bonner's salts, 250 $\mu\text{g ml}^{-1}$ streptomycin with or without 5 mM cyclic AMP, in a 5 cm Petri dish. Washed, sterile Cellophane (slit Visking dialysis tubing) held in perspex stretchers (each consisting of an inner ring fitting tightly into an outer one) was placed over the cells and air bubbles excluded. At the stated times of starvation the central area of agar and Cellophane was cut out and the cells fixed *in situ* with ice-cold fresh fixative (10 mM sodium-cacodylate, 0.5% OsO_4 , 0.5% glutaraldehyde, 1 mM MgCl_2 , 1 mM CaCl_2 , pH 6.8). After 20–30 min the cells were collected, washed three times with 10 mM sodium cacodylate, pH 6.8 and dehydrated in graded alcohols. The specimens were then infiltrated in acetone for 1 h, then in a 1:1 mixture of acetone and Spurr's resin for at least 6 h and finally polymerised in Spurr's resin at 60 °C for 24 h. Sections were stained with lead citrate and uranyl acetate and photographs taken of random fields of cells with a Siemens Elmiskop 1 electron microscope. Prespore vacuoles were seen in 28%, 37%, 26%, 2% and 4% of the cell profiles at 15, 20, 26, 37 and 52 h respectively of induction with cyclic AMP.

Fig. 2 Induction of prespore and stalk cells by cyclic AMP at low (a) or high (b) cell density. ○, Prespore cells, no cyclic AMP; □, stalk cells, no cyclic AMP; ●, prespore cells + cyclic AMP; ■, stalk cells + cyclic AMP. HMI cells were prepared and plated as described in Fig. 1. Those at high density ($2.5 \times 10^5 \text{ cm}^{-2}$) were overlaid with Cellophane but those at low ($5 \times 10^2 \text{ cm}^{-2}$) were not. At the stated times cells at high density were collected in KK_2 and disaggregated with Pronase/British anti-lewisite (BAL)¹³. Those at low density were detached by swirling in KK_2 lacking Mg^{2+} and then pelleted in siliconised centrifuge tubes for 10 min at 1,000g. In both cases the cells were then fixed in 60% methanol and fluorescently stained using absorbed anti-spore serum by a modification¹⁴ of the method of Hayashi and Takeuchi¹³. Cells were examined in a Zeiss Photomicroscope III equipped for epi-fluorescence.



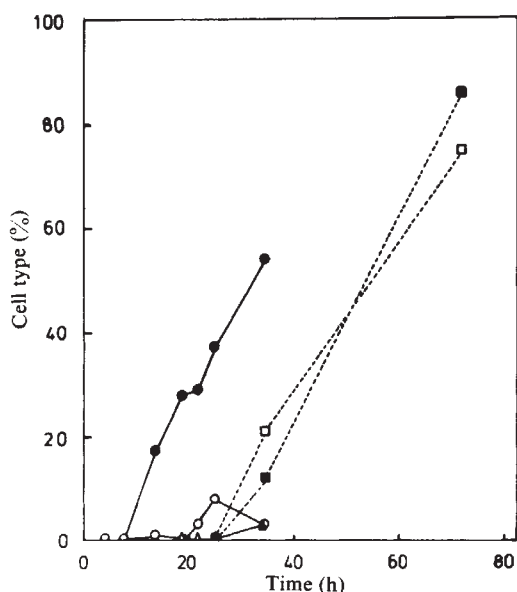


Fig. 3 Stalk and lack of prespore cell differentiation in cells at low density adjacent to high density 'helper' cells. ○, Prespore cells in low density cells over 'helper'; △, prespore cells in low density cells alone; □, stalk cells in low density cells over 'helper'; ●, prespore cells in 'helper'; ■, stalk cells in 'helper'. HMI cells were collected and plated on 5 mM cyclic AMP agar at a density of $2.5 \times 10^5 \text{ cm}^{-2}$ as described in Fig. 1. These 'helper' cells were overlaid with thin Cellophane (325 P, British Cellophane) held in a stretcher. More cells at low density (10^3 cm^{-2}) were then spread over the Cellophane. In control plates the 'helper' cells were omitted. At various times a cover slip was placed on the upper layer of cells and stalk cells scored above and below the Cellophane by microscopy. The low density cells were then detached from the Cellophane by swirling for 10 min with 50 mM Tris, 0.1% Pronase, 25 mM BAL pH 7.0¹³ and pelleted in siliconised centrifuge tubes at 1,000g for 10 min. 'Helper' cells were collected and both sets of cells stained as described in Fig. 2. The recovery of low density cells was normally 5–30% of those plated. Several controls indicate that there was not a selective loss of prespore cells during cell collection. First, similar results were obtained when cells were detached using the Pronase/BAL mixture diluted 1:10 or when they were scraped off in KK_2 and when they were pelleted at $\sim 10,000g$ for 30 s in a microfuge. Second, disaggregated high density cells, plated at low density could be recovered after 1 h with an efficiency of 14–50%. The losses occurring during cell collection did not in this case significantly change the proportion of prespore cells.

many must eventually transform into stalk cells. Support for this view comes from electron microscopic observations of cells taken at later times which show that many PSVs are in the process of breaking down, in a way resembling that described as occurring during the 'de-differentiation' of isolated prespore cells¹⁷. Such a transformation is also consistent with the observations of Town

and Stanford¹⁸ that prespore cells taken from slugs produce a high proportion of stalk cells (but no spores) in similar conditions.

As stalk cells are not inducible by cyclic AMP at low cell density we wondered if the same would be true for prespore cells. Figure 2a shows that this is indeed the case. From this result we conclude that a cell interaction must also be involved in the induction of prespore cells. We showed previously that stalk cell formation by cells at low density ('tester cells') was stimulated by the presence of a high density layer of cells ('helper cells') separated from them by a Cellophane membrane⁷. It was of interest therefore to see if the same was true for the formation of prespore cells. The results of a 'helper' experiment of this type, in which we assayed for stalk and prespore cells, is shown in Fig. 3. Stalk cells are induced in the helper and tester cells with near identical efficiency and kinetics. This is not true for prespore cells. Although a slight induction of prespore cells is consistently observed in the tester cells (compared to low density controls) it is much lower than in the helper cells themselves. Evidently the requirements for induction of prespore cells differs in some way from those for stalk cells.

In the light of these results it is apparent that the failure of wild-type amoebae to form spores when incubated with cyclic AMP under Cellophane at high cell density is not due to their failure to enter the prespore pathway but rather to some block in the maturation of prespore cells. Mutants can be isolated that do form spores in such conditions⁷. These mutants have presumably overcome the unknown block to spore maturation but might be expected to show the same requirement as wild-type amoebae for entry into the prespore pathway. We have, therefore, examined the ability of low density preparations of such strains to be helped to differentiate into stalk and spore cells by high density cells separated from them by a sheet of Cellophane. At the densities used the low density tester cells gave on their own 0–3% stalk cells and no spores. The results for five mutant strains presented in Table 1 show that as expected each is efficiently helped to make stalk cells but not to make spores.

From Table 1 it is evident that there is some stimulation of spore formation by the helper cells, especially when the density of the tester cells was $5 \times 10^3 \text{ cm}^{-2}$, the same density as used previously⁷. We have found (data not shown) that the efficiency of helping is not improved by using Cellophane treated with zinc chloride to increase its porosity. Further we have noted consistently that the spores formed as a result of helping are almost invariably very close to other cells, whereas most of the stalk cells are isolated (Table 1). Both these observations indicate that one of the requirements for entry into the prespore pathway involves a short-range cellular interaction, not required for stalk cell formation. The small stimulation of spore formation by helper cells is consistent with the idea that a low-molecular weight factor is required for both pathways of differentiation.

In conclusion, our results show that in appropriate conditions exogenous cyclic AMP can induce the formation of stalk cells and also the formation of prespore cells or even of mature spores.

Table 1 Enhancement of stalk and spore cell differentiation in low density mutant cells by the same cells at high density

Strain	% Differentiated cell type observed					
	Above cellophane $10^3 \text{ cells cm}^{-2}$			Below cellophane $2.5 \times 10^5 \text{ cells cm}^{-2}$		
	Total	Stalk Isolated (as % of stalk)	Spore Isolated (as % of spore)	Total	Stalk	Spore
HM9	15.5	90	0	11.9	26.7	
HM9*	18.0	49	1.3	14.6	22.2	
HM13	36.5	79	0.5	21.5	37.4	
HM15	18.2	93	0	27.0	22.3	
HM16	15.6	68	0	24.5	8.1	
HM18	45.7	65	3.3	44.5	29	

Cells of the mutants were collected as described in Fig. 1. High density cells were plated directly onto 5 mM cyclic AMP agar and overlaid with thin Cellophane as in Fig. 3. Cells of the same strain were spread on top at low density. After 3 d incubation at 22 °C a cover slip was placed on the top layer of cells and the cells scored above and below the Cellophane by microscopy. Strains HM13, 15, 16 and 18 were obtained by mutagenesis and selection from HM2 (a cobalt resistant strain temperature-sensitive for growth derived from HMI; R.R.K. & J. Trent, unpublished). Strain HM9 is clone 27 of the strain previously called *sci-1*.

*HM9 cells above Cellophane at a density of $5 \times 10^3 \text{ cm}^{-2}$.

Since cells under Cellophane become aggregation competent in the absence of cyclic AMP (R.R.K., unpublished observations) there must be a requirement for cyclic AMP in both pathways of differentiation beyond the stage of aggregation competence. We therefore regard with caution the idea that levels of cyclic AMP control the pattern of differentiation in *Dictyostelium*^{19,20}. We have, however, identified different requirements for the induction of stalk cells and of prespore cells (in the wild type) or spores (in the mutants). These differences could play a key part in controlling the pattern of differentiation in normal development.

We thank Dr David Forman for producing the antibody, James Vernon and Jennifer Trent for assistance and Dr Julian Gross for helpful discussions.

Imperial Cancer Research Fund, UK

University of Southampton,
Department of Biology,
Medical and Biological Sciences Building,
Bassett Crescent East,
Southampton, UK

Imperial Cancer Research Fund,
Mill Hill Laboratories,
Burtonhole Lane,
London NW7, UK

R. R. KAY

D. GARROD

R. TILLY

Received 15 August; accepted 2 November 1977.

1. Kauffman, S. *Ciba Fdn Symp.* **29**, 210-222 (1975).
2. Gardner, R. L. & Rossant, J. *Ciba Fdn Symp.* **40**, 5-25 (1975).
3. Wolpert, L. *Curr. Top. dev. Biol.* **6**, 183-224 (1971).
4. Raper, K. B. J. *Elisla Mitchell Scientific Soc.* **56**, 241-282 (1940).
5. Bonner, J. T. *Proc. natn. Acad. Sci. U.S.A.* **65**, 110-113 (1970).
6. Francis, D. *Nature* **258**, 763-765 (1975).
7. Town, C. D., Gross, J. D. & Kay, R. R. *Nature* **262**, 717-719 (1976).
8. Hohl, H. R. & Hamamoto, S. T. *J. Ultrastruct. Res.* **26**, 442-453 (1969).
9. Maeda, Y. & Takeuchi, I. *Development Growth Differentiation* **11**, 232-245 (1969).
10. Gregg, J. H. & Badman, S. *Dev. Biol.* **22**, 96-111 (1970).
11. Müller, U. & Hohl, H. R. *Differentiation* **1**, 267-276 (1973).
12. Takeuchi, I. *Dev. Biol.* **8**, 1-26 (1963).
13. Hayashi, M. & Takeuchi, I. *Dev. Biol.* **50**, 302-309 (1976).
14. Forman, D. & Garrod, D. R. *J. Embryol. exp. Morph.* **40**, 215-228 (1977).
15. Yu, N. Y. & Gregg, J. H. *Dev. Biol.* **47**, 310-318 (1975).
16. Takeuchi, I. & Sakai, Y. *Development Growth Differentiation* **13**, 201-210 (1971).
17. Sakai, Y. & Takeuchi, I. *Development Growth Differentiation* **13**, 231-240 (1971).
18. Town, C. D. & Stanford, E. J. *Embryol. exp. Morph.* (in the press).
19. Pan, P., Bonner, J. T., Wedner, H. J. & Parker, C. W. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1623-1625 (1974).
20. McMahon, D. *Proc. natn. Acad. Sci. U.S.A.* **70**, 2396-2400 (1973).

Unintegrated ribosomal genes and their relation to position effect variegation in *Drosophila melanogaster*

THE variegated expression of genes is most often associated with a chromosomal rearrangement having one or both breakpoints within a heterochromatic region¹⁻³ and as a result, the genes lying near the breakpoint are usually suppressed. In *Drosophila melanogaster*, the genes coding for ribosomal RNA (rRNA) at the nucleolus organiser region are also situated in heterochromatin⁴. Ribosomal genes which are not integrated into the DNA of the chromosome have been found in *D. melanogaster* females heterozygous for an X chromosome inversion which transposes the nucleolus organiser to the tip of the X (refs 5, 6). It was proposed that these unintegrated ribosomal genes ($\sim 3 \times 10^8$ daltons) might be involved in position effect variegation⁵. This study was initiated to determine if a correlation exists between the presence of unintegrated ribosomal genes and the extent of position effect variegation in general. DNA from inversion-bearing male genotypes undergoing position-effect variegation was examined by sucrose gradient sedimentation and DNA-rRNA hybridisation. The resulting data further support a variegation association.

When the nucleolus organiser is moved by an inversion to the tip of the X chromosome, it has been shown to variegates⁷. Baker⁷ inferred from the lethality associated with *In(1)sc^{L8}/0* and *In(1)sc^{S1}/0* males that these two X chromosome inversions were associated with a position effect suppression of ribosome synthesis. In the first series of experiments, the X chromosome

inversions used moved the nucleolus organiser to the tip of the X (*In(1)sc^{V2}*, *sc^{L8}*, *sc^{L8}*, and *sc^{S1}*). In (*1*)*sc^{L8}*, on the other hand, has its proximal break distal to the nucleolus organiser and therefore left it at the base of the X. Oregon-R X chromosome was used as a wild-type control. DNA with a molecular weight of $\sim 5 \times 10^8$ was isolated from diploid brain and imaginal disks, and polytene salivary glands of third instar male larvae. Each male carried a specific X inversion and lacked a Y chromosome (*In(1)X/0*). As shown in Table 1, Oregon-R X/0 and *In(1)sc^{L8}/0* males have no low molecular weight ribosomal genes in diploid tissue. rRNA hybridisation was within the high molecular weight DNA peak. Diploid larval tissue was analysed because of the lethality associated with some of the *In(1)X/0* genotypes⁷. The pattern of ribosomal gene integration found for adult flies has always been identical to that of the larval diploid tissue^{5,8,9}. All male genotypes which carry an inverted X with the right breakpoint between the nucleolus organiser and the centromere (*In(1)sc^{V2}*, *sc^{L8}*, *sc^{L8}*, and *sc^{S1}*) (Table 1) exhibit unintegrated ribosomal genes ($\sim 3 \times 10^8$ daltons) in increasing amounts. Of particular interest is the fact that *In(1)sc^{L8}/0* and *In(1)sc^{S1}/0* males produce the highest percentage of unintegrated ribosomal genes in diploid tissue and they also suffer the greatest lethality, since few males were observed to survive to third larval instar.

When the DNA from the polytene salivary glands was examined, all male genotypes had approximately 42% unintegrated ribosomal genes (Table 1). This percentage of low molecular weight genes in the salivary glands never varies, regardless of the genotype, and the finding has been discussed previously^{5,8}.

Mechanisms operating at the chromosomal level to affect gene expression may also be enhanced or suppressed by using suppressors of position-effect variegation³. One such locus is suppressor of variegation (*Su(var)*) in *D. melanogaster*. *Su(var)* is not a general modifier of position effect, it is limited to rearrangements involving specific heterochromatic regions (for example, *In(1)w^{m4}* and *rst^a* compared with *In(1)sc^{L8}*) (refs 10-13). In the second series of experiments, high molecular weight DNA was isolated from those inversion-bearing adult male genotypes which are affected by the *Su(var)* locus. The studies reported here were conducted using two alleles, *Su(var)* and *Su(var)⁺*, which can be maintained in homozygous stocks. *Su(var)⁺* enhances the variegation associated with *In(1)w^{m4}* and *In(1)rst^a*, whereas *Su(var)* enhances variegation in *In(1)sc^{L8}* (refs 3, 12, 14). When these suppressor alleles were associated with X chromosome inversions which enhance variegation, unintegrated ribosomal genes were present in all of the adult male genotypes analysed (Table 2). When variegation was suppressed by using *Su(var)* in association with *In(1)w^{m4}* and *In(1)rst^a* and *Su(var)⁺* with *In(1)sc^{L8}* (refs 3, 12, 14), no unintegrated ribosomal genes were found (Table 2).

Table 1 Percentage of unintegrated rDNA in male *D. melanogaster* of different genotypes

Genotype	Diploid brain and imaginal disks	Polytene salivary glands	No. of gradients analysed for each tissue
Ore-R X/0	0	42.1	3
<i>In(1)sc^{L8}/0</i>	0	41.9	3
<i>In(1)sc^{V2}/0</i>	5.2	41.1	3
<i>In(1)sc^{S1}/0</i>	9.6	42.3	3
<i>In(1)sc^{L8}/0</i>	39.1	42.6	3
<i>In(1)sc^{S1}/0</i>	46.2	41.7	3

Drosophila melanogaster were labelled by growth at 23-25 °C on a low yeast medium as described previously^{5,8}. The medium contained 5 μ Ci ml⁻¹ ³H-methyl thymidine (50-60 mCi mmol, New England Nuclear). DNA isolation, sedimentation, and hybridisation of ³H-rRNA to the gradient fractions were as before^{5,8}. The *Drosophila* tissue culture rRNA used for the hybridisation had a specific activity of 309,000 c.p.m. μ g⁻¹. The stocks and inversions used in this study have been described previously^{5,8}. For genetic symbols and descriptions, refer to Lindsley and Grell¹⁹.

Table 2 Percentage of unintegrated rDNA in progeny adult males from female *D. melanogaster* of differing phenotype following inversion

Inversion	Mother	Progeny adult males	rDNA Unintegrated (%)
<i>w^{m1}</i>	Su(var) ⁺ /Su(var) ⁺	Su(var) ⁺ /Su(var) ⁺	17.8
<i>w^{m4}</i>	Su(var)/Su(var)	Su(var)/Su(var)	0
<i>rst³</i>	Su(var) ⁺ /Su(var) ⁺	Su(var) ⁺ /Su(var) ⁺	16.2
<i>rst³</i>	Su(var)/Su(var)	Su(var)/Su(var)	0
<i>sc¹</i>	Su(var)/Su(var)	Su(var)/Su(var)	18.2
<i>sc¹</i>	Su(var) ⁺ /Su(var) ⁺	Su(var) ⁺ /Su(var) ⁺	0

The *Drosophila* stocks and X chromosome inversions used have been described previously^{3,11,14}. The construction of genotypes was essentially the method of Spofford^{11,13}. Each inversion was placed opposite an attached X chromosome carrying the markers *y* (yellow) and *w* (white) and then introduced into companion stocks. The companion stocks were homozygous for *Su(var)* or for the *Su(var)⁺* allele. This was done by a series of three crosses to replace the third chromosome of the inversion stock with *Ly Sb (Lyra, Stubble)* chromosome and then replace the latter with the *Su(var)* or *Su(var)⁺* chromosome from coisogenic stocks. (The stocks carrying *Su(var)* or *Su(var)⁺* were a gift of W. K. Baker, Chicago.) *Su(var)* and *Su(var)⁺* females used for the crosses were tested for the *Su(var)* locus genotype by mating to Dp(1;3)w^{m261-384} males^{10,12}. For all the inversions studied, crosses between *Su(var)⁺/Su(var)⁺* or *Su(var)/Su(var)* were set up on ¹⁴C-methyl thymidine-labelled food. Adult sons carrying *Su(var)* or *Su(var)⁺* and the inversion *ln(1)w^{m4}, rst³, or sc¹* were picked from the progeny of each mating pair and their DNA was subjected to sucrose gradient sedimentation and hybridization to ³H-*Drosophila* rRNA. Three gradients were analysed for each inversion.

It seems that there is a diverse type of regulation of nucleolus organiser activity in *D. melanogaster*. This regulation can be expressed genetically as a variegated suppression of nucleolus organiser activity, or expressed physically in the form of unintegrated ribosomal genes. It is known that the physical structure of ribosomal DNA not only differs from the bulk of the DNA (ref. 15) but it also differs within the ribosomal DNA itself depending on the genotype^{5,6,8,9}. In *D. melanogaster* the nucleolus organiser, together with an undetermined extent of heterochromatin to either side of it, can also induce variegation in its neighbours when placed into euchromatin, for example near *ct*, *l*, or *in* (ref. 16). Inversions such as *ln(1)sc^{V2}*, *sc⁸*, *sc^{L8}* and *sc^{S1}* move a large amount of heterochromatin to the tip of the X chromosome, a condition which could favour variegation. Interestingly, Baker⁷ reported a decreasing viability induced by suppression of the nucleolus organiser for *ln(1)X/0* male genotypes (*sc¹/0:95*; *sc^{V2}/0:38*; *sc⁸/0:14* or *<1*; and *sc^{S1}/0:<1*). In this study, I have found an increasing amount of unintegrated ribosomal genes as the apparent male viability decreased compared with female survival, and as the position effect increased. The viability of *ln(1)sc^{L8}/0* and *ln(1)sc^{S1}/0* males is partially restored by Y chromosomes with an altered nucleolus organiser and fully restored by unaltered Y chromosomes⁷. When these inversions were analysed in the presence of a Y, no unintegrated ribosomal genes were found^{5,6}.

ln(1)sc^{S1}/0 male larvae were reported to have a 15% reduction in the amount of 18S and 28S rRNA (refs 17, 18). It was suggested that this reduction results from a suppression of transcription: perhaps it is the unintegrated ribosomal genes which cannot be transcribed efficiently. Furthermore, it has been suggested that the *Su(var)* locus might also affect a signal to begin message transcription¹⁴. It is not apparent, however, why the *Su(var)* locus should affect the nucleolus organiser region. The purpose of unintegrated ribosomal genes is still not understood. These genes have been postulated to be a result of chromosome pairing in inversion-bearing heterozygous females⁶. They have never been observed in adult males or in diploid tissue from male genotypes. The data presented here are the first to show that unintegrated ribosomal genes can occur in male genotypes of *D. melanogaster*. This report has also

presented two experimental situations in which position-effect variegation affects particular X chromosome inversions and also results in the presence of low molecular weight ribosomal genes. Little is known of the mechanism of position-effect variegation, but by studying chromosome rearrangements and factors which influence gene expression, a molecular explanation for the role of unintegrated genes in the variegation process may be found.

I thank Drs A. Gayler Harford and Thomas G. Sanders for reading the manuscript and for their helpful comments.

CHRISTINE I. ZUCHOWSKI-BERG

Department of Biology,
Princeton University,
Princeton, New Jersey 08540

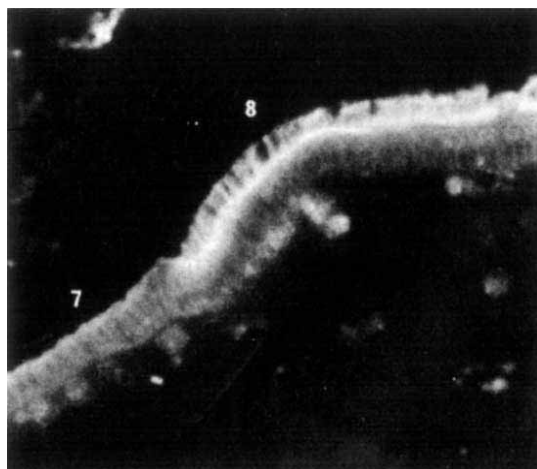
Received 15 August; accepted 1 November 1977.

1. Lewis, E. B. *Adv. Genet.* **3**, 73-115 (1950).
2. Baker, W. K. *Adv. Genet.* **14**, 133-169 (1968).
3. Spofford, J. B. *The Genetics and Biology of Drosophila 1c* (eds W. Ashburner & X. Novitski), 955-1018 (Academic London, 1976).
4. Ritossa, F. M. & Spiegelman, S. *Proc. natn. Acad. Sci. U.S.A.* **53**, 737-745 (1965).
5. Zuchowski, C. I. & Harford, A. G. *Cell* **11**, 383-388 (1977).
6. Harford, A. G. & Zuchowski, C. I. *Cell* **11**, 389-394 (1977).
7. Baker, W. K. *Proc. natn. Acad. Sci. U.S.A.* **68**, 2472-2476 (1971).
8. Zuchowski, C. I. & Harford, A. G. *Chromosoma (Berl.)* **58**, 219-234 (1976).
9. Zuchowski, C. I. & Harford, A. G. *Chromosoma (Berl.)* **58**, 235-246 (1976).
10. Cohen, J. *Genetics* **47**, 647-659 (1962).
11. Spofford, J. B. *Genetics* **57**, 751-766 (1967).
12. Spofford, J. B. *Drosophila Inf. Serv.* **50**, 98 (1973).
13. Spofford, J. B. *Proc. 11th Int. Congr. Genet., The Hague*, **1**, 41 (1965).
14. Spofford, J. B. *Genetics* **62**, 555-571 (1969).
15. Harford, A. G. & Zuchowski, C. I. *Biochem. biophys. Res. Commun.* (in the press).
16. Hannah-Alava, A. *Molec. gen. Genet.* **113**, 191-203 (1971).
17. Nix, C. E. *Biochem. Genet.* **10**, 1-12 (1973).
18. Puckett, L. D. & Snyder, L. A. *Exp. Cell Res.* **95**, 31-38 (1975).
19. Lindsley, D. L. & Grell, E. H. *Carnegie Inst. Wash. Publ.* 627 (1968).

Evidence for a 'mammalian' thyroglobulin in endostyle of the ascidian *Styela clava*

ATTENTION has focused on the protochordate endostyle since the demonstration that certain of its cells could bind iodine¹⁻⁴. That has been confirmed by ultrastructural studies⁴⁻⁸, while the presence of thyroid hormones and their precursors has been established in endostylar extracts⁹⁻¹⁰. These findings have been questioned because the blood¹¹ and tunic^{9,13} can bind large amounts of iodine. Moreover, both iodotyrosines and iodothyronines can be isolated from tunic extracts^{9,12-15}. Barrington has suggested that iodine-binding by the tunic is an invertebrate process associated with the extracellular iodination of inert structural scleroproteins, whereas endostylar iodination involves intracellular glycoproteins⁸. Using indirect immunofluorescence¹⁶, with rabbit anti-thyroglobulin serum, I have now identified a

Fig. 1 Photomicrograph of zones 7 and 8 from the endostyle of *S. clava*. Note the brilliant apical fluorescence in the ciliated zone 8. The fluorescence at the tips of zone 7 cells rather less distinct. Zeiss epi-fluorescence, mercury vapour lamp, blue excitation ($\times 300$).



protein in the endostyle of the ascidian *Styela clava* that is immunologically identical to mammalian thyroglobulin.

Figure 1 shows the specific localisation of thyroglobulin in zones 7 and 8 of the endostyle of *S. clava*. The distinctly brighter fluorescence in zone 8 may reflect fundamental differences in the quantity and/or quality of the intracellular material. Equally it may simply be due to secretory material trapped at the bases of cilia. Routine controls using saline or non-immune serum proved negative.

Thus, in the regions of the endostyle that are known actively to bind iodine intracellularly, the protein to which it is evidently bound is immunologically indistinguishable from mammalian thyroglobulin. Unquestionably protochordates must be regarded as very specialised animals¹⁷, and the extent to which the iodine-binding cells themselves have also become specialised is difficult to ascertain.

Notwithstanding such difficulties, this latest evidence for a thyroglobulin-like glycoprotein in the iodine-binding cells, together with the earlier work cited here clearly makes the concept of the endostylar origin of vertebrate thyroidal cells more credible.

MICHAEL C. THORNDYKE

Department of Zoology,
Bedford College,
Regent's Park,
London NW1, UK

Received 29 July; accepted 21 October 1977.

1. Barrington, E. J. W. & Franchi, L. L. *Nature* **177**, 432 (1956).
2. Thomas, I. M. *J. mar. biol. Ass. U.K.* **35**, 203-210 (1956).
3. Barrington, E. J. W. *J. mar. biol. Ass. U.K.* **36**, 1-16 (1957).
4. Barrington, E. J. W. & Thorpe, A. *Gen. comp. Endocr.* **5**, 373-385 (1965).
5. Fujita, H. & Nanba, H. *Z. Zellforsch. mikrosk. Anat.* **121**, 455-469 (1971).
6. Thorpe, A., Thorndyke, M. C. & Barrington, E. J. W. *Gen. comp. Endocr.* **19**, 559-571 (1972).
7. Dunn, A. D. *J. exp. Zool.* **188**, 103-123 (1974).
8. Thorpe, A. & Thorndyke, M. C. *Symp. zool. Soc. Lond.* **36**, 159-177 (1975).
9. Roche, J., Salvatore, G. & Rametta, G. *Biochim. biophys. Acta* **63**, 154-165 (1962).
10. Barrington, E. J. W. & Thorpe, A. *Proc. R. Soc., Lond.* **B163**, 136-149 (1965).
11. Kennedy, G. R. *Gen. comp. Endocr.* **7**, 500-511 (1966).
12. Barrington, E. J. W. *Symp. zool. Soc. Lond.* **36**, 129-158 (1975).
13. Salvatore, G., Vecchio, G. & Macchia, V. *C.R. Seanc. Soc. Biol.* **154**, 1380-1384 (1960).
14. Roche, J., Rametta, G. & Varrone, S. *Gen. comp. Endocr.* **4**, 277-284 (1962).
15. Barrington, E. J. W. & Thorpe, A. *Proc. R. Soc., Lond.* **B171**, 91-109 (1968).
16. Coons, A. H. & Kaplan, M. H. *J. exp. Med.* **91**, 1-13 (1950).
17. Barrington, E. J. W. *The Biology of Hemichordata and Protochordata* (Oliver & Boyd, Edinburgh, 1965).

Hormonal influences on ocular refraction in the rat

We report here that thyrotrophin (TSH) has a marked influence on ocular refraction in rats. In the clinical experience of one of us (C.C.K.-V.A.), adolescent myopia is not infrequently accompanied by increased height, suggesting a common, possibly hormonal, cause for both phenomena. No data on possible hormonal influences on eye growth or refraction could be found in the literature, but early anatomical studies showed that, during puberty, a growth spurt occurs in the eye after an early post-natal rapid weight increase¹⁻⁴. In a very detailed study, Weiss² found relatively less growth of the sagittal eye axis as compared with the vertical and frontal diameters. Non-uniform growth rates along different eye axes might affect refraction.

In an attempt to assess hypophyseal influences on the eye, we studied the effects of some pituitary hormones on refraction and eye growth in rats. Seven-week-old ('adolescent') male Wistar rats were hypophysectomised and treated with a pituitary TSH preparation (Ambion, Organon), which was dissolved in saline and injected subcutaneously twice daily, starting on the day after hypophysectomy. From preliminary experiments, a treatment period of 3 weeks was found to be optimal and was subsequently used unless indicated otherwise. Control rats were injected with saline alone. Ocular refraction was measured to the nearest dioptre by direct ophthalmoscopy after accommodation had been paralysed by a drop of 1%

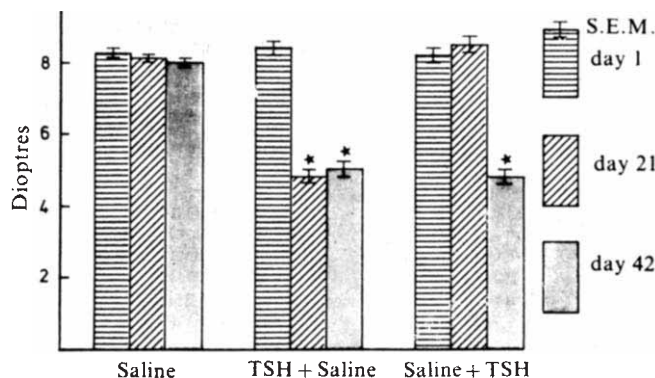


Fig. 1 Irreversibility of TSH effects on ocular refraction in hypophysectomised rats. Hypophysectomy was on day 1 and injections from day 2, twice daily 0.1 ml subcutaneously. Controls received saline from day 2 to day 42 ($n = 9$). TSH + saline: 2 IU TSH d^{-1} from day 2 to day 21, saline from day 22 to day 42 ($n = 9$). Saline + TSH: saline from day 2 to day 21, 2 IU TSH d^{-1} from day 22 to day 42 ($n = 8$). Autopsy was carried out on day 42. Averages of refraction in both eyes per rat were used for computation. Columns represent means $\pm$ s.e.m. * $P < 0.01$, compared with intact controls, computed by analysis of variance.

atropine solution, 1 h before examination. The animal was restrained manually and the entrance of the superior temporal vein in the ocular fundus was brought into focus by interposing the appropriate ophthalmoscope lens. In Tables 1 and 2, 'refraction' refers to the strength of this compensating lens, expressed in dioptres. It should be stressed that the high positive values of the 'refraction' do not mean that rats have a physiological hypermetropia, as it is known that small-eyed animals tend to exhibit large hypermetropia^{5,6} attributable to a measurement artefact arising from the ophthalmoscopic reflection occurring in front of the receptors⁷. The eyes of most animals are, in fact, thought to be emmetropic. The term 'hypermetropia' is used here, however, for the sake of brevity. Ophthalmoscopy was always carried out by the same investigator, using an eye that could not accommodate. In the experiment of Table 1, the ophthalmoscopist did not know to which group an animal belonged. An unbiased adherence to such a 'blind' procedure was not always possible in other experiments, as operation scars, growth retardation and behavioural changes were unavoidable clues as to an animal's treatment. (In frequent 'blind' checks, however, the reproducibility of measurements was ascertained.) Refraction changes due to stress-induced accommodation, such as occurs in teleosts⁷, could be excluded since the rats were cycloplegic after atropinisation. One day after the last ophthalmoscopic examination, the animals were killed with chloroform. The eyes were enucleated and weighed on a torsion balance; the lenses were then extracted and also weighed.

The results in Table 1 show that TSH brings about a dose-dependent change in refraction. No significant changes in eye or lens weights were evident.

In order to assess whether endogenous TSH has comparable effects, its production was enhanced by thyroidectomy on 7-week-old rats in a subsequent experiment. As is shown in Table 2, 3 weeks later, refraction had changed markedly compared with intact controls, although eye and lens weights were unchanged. In contrast, control animals which had only been hypophysectomised, had markedly lower eye and lens weights, but their refraction remained normal. The results of this experiment were taken as evidence for a direct effect of TSH on the eye, also occurring in the absence of the thyroid gland.

We then investigated whether the ocular effect of TSH is reversible. One group of rats hypophysectomised at 7

Table 1 Effect of exogenous thyrotrophin on ocular refraction and eye weight in hypophysectomised rats

Treatment	n	Body weight (g)		Refraction (D)				Weight total 2 eyes (mg)	Weight total 2 lenses (mg)
		Day 1	Day 21	Day 1		Day 21			
				L	R	L	R		
Saline	11	201	190	8±0.6	8±0.8	8±0.6	8±0.5	216±6	62±3.8
TSH 1 IU d ⁻¹	11	200	202	8±0.8	8±0.8	5±0.7*	5±0.5*	221±7	62±4.7
TSH 2 IU d ⁻¹	11	199	210	8±0.3	8±0.0	4±0.8†	4±0.6†	222±6	64±4.7

In all experiments, rats were exactly 7-weeks-old on day 1, when hypophysectomy was performed. Injections of TSH or saline (0.1 ml subcutaneously twice daily) were made from day 2 until killing on day 21. D: Refraction (in dioptres) measured by ophthalmoscopy. Analysis of variance:

* Statistically significant, $P < 0.01$, from control values and from TSH 2 IU d⁻¹.

† $P < 0.01$, from control values and from TSH 1 IU d⁻¹. Means ± standard deviation.

weeks of age was treated with 2 IU TSH daily for 3 weeks and subsequently with saline for 3 weeks. In a second group, a crossover design was followed, whereby the animals received saline for 3 weeks, followed by 2 IU TSH d⁻¹ for 3 weeks. Hypophysectomised controls received saline for 6 weeks. At the end of the experiment, both groups treated with TSH showed the same change in refraction as seen before (Fig. 1). The animals treated first with TSH, then with saline, showed the expected refraction

were treated with either 2 IU TSH d⁻¹, or ACTH (1 U per 100 g body weight subcutaneously every alternate day), or a combination of the two hormones in the same dosages; controls received saline. Treatment lasted for 3 weeks and the results are shown in Fig. 2. TSH decreased hypermetropia: ACTH when given alone had no effect. When the two hormones were given simultaneously, however, the effect of TSH was not only blocked but even reversed, resulting in increased hypermetropia.

Table 2 Effect of exogenous thyrotrophin on ocular refraction and eye weight in hypophysectomised rats

Treatment	n	Body weight (g)		Refraction (D)				Weight total 2 eyes (mg)	Weight total 2 lenses (mg)
		Day 1	Day 21	Day 1		Day 21			
				L	R	L	R		
Intact	10	190	309	8±0.7	8±0.3	8±0.4	8±0.4	226±7	62±2.3
Hypophysectomised	13	189	173	8±0.9	8±1.0	8±0.4	8±0.5	212±7*	58±3.4*
Thyroidectomised	10	190	251	8±0.5	8±0.8	5±0.5*	5±0.5*	227±6	59±6.7

Either hypophysectomy or thyroidectomy was performed on days 1 and 2. Animals were left untreated until killing on day 21. Analysis of variance: * Statistically significant, $P < 0.01$, from control values.

alteration, indicating that these effects persist for at least 3 weeks after discontinuation of TSH treatment.

From Table 2 it follows that in intact animals, in spite of the presumably continuous presence of TSH, no decrease in hypermetropia occurs between 7 and 10 weeks of age. Obviously, in such rats, compensatory influences are operative. In a first analysis of such a compensation by other pituitary hormones, 7-week-old hypophysectomised rats

In rats, the Harderian gland is the major retrobulbar constituent of orbital contents, having about the same weight as the eye⁸. Changes in gland size might cause deformation of the eye bulb and hence refraction changes. Therefore we looked for parallelism between changes in Harderian gland size and refraction under the influence of exogenous or endogenous TSH. Three weeks after hypophysectomy there was a marked atrophy of the Harderian gland, but refraction remained normal (Fig. 3). After TSH treatment of hypophysectomised animals, the decrease in hypermetropia was not accompanied by a change in gland atrophy. The rise of endogenous TSH after thyroidectomy, however, does not affect Harderian gland size, it only changes refraction. We therefore conclude that there is no causal relationship between Harderian gland size and refraction.

In conclusion, TSH induces changes in refraction of the rat eye by a process involving neither thyroid hormone nor the Harderian gland. Experiments to elucidate the physical, morphological and biochemical mechanisms causing refraction changes are in progress, with special attention being paid to corneal radius and refraction, and lenticular refraction and biochemistry.

No literature data have come to our knowledge on modifications of ocular refraction in the rat. As to eye growth in the rat, it is known that after very rapid growth in early life, the weight increases at a slower rate during the remaining lifespan⁹⁻¹¹. Although some factors such as food, care and physical condition have been shown to influence relative and absolute eye weight⁶, no specific hormonal factors have until now been held responsible for such phenomena. If TSH can be shown to affect refraction in man in a similar way to that in the rat, the data presented here open up

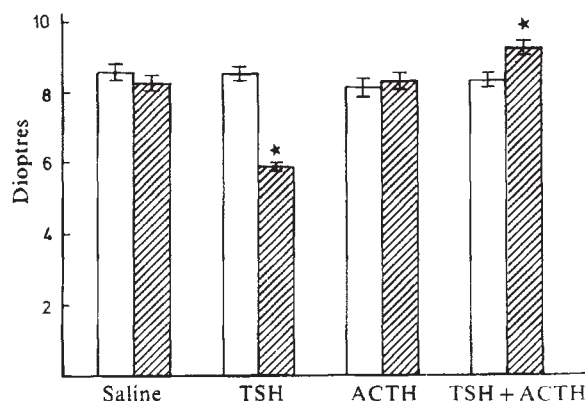


Fig. 2 Blockade by ACTH of TSH effect on ocular refraction in hypophysectomised rats. Hypophysectomy was on day 1. TSH (2 IU d⁻¹) and saline were injected twice daily. ACTH (1 IU per 100 g body weight) was injected every other day. No significant differences in eye weights were found. * $P < 0.01$, compared with saline-treated controls, computed by analysis of variance. Six animals in each group. Open columns, day 1; hatched columns, day 21.

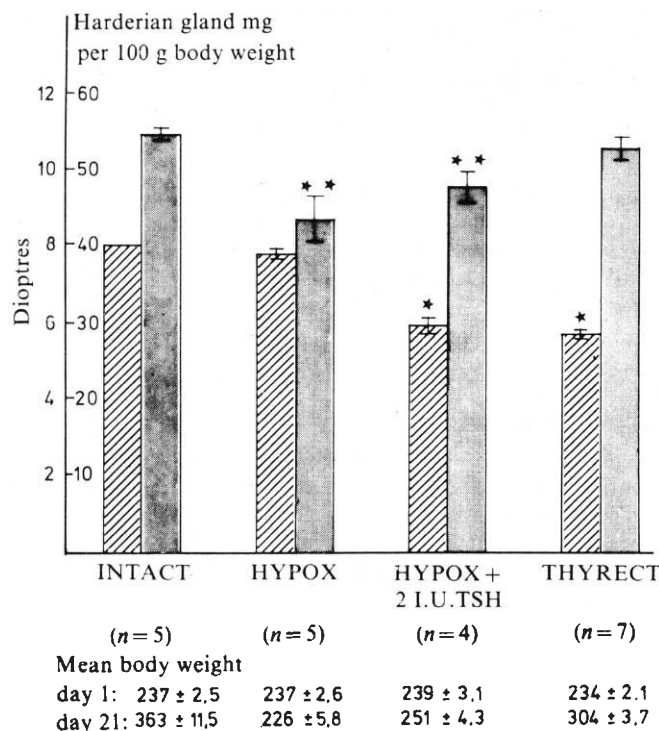


Fig. 3 Effect of exogenous and endogenous TSH on Harderian glands and refraction in rats. Hypophysectomy (hypox.) or thyroidectomy (thyrect.) was performed on day 1. TSH treatment (2 IU d^{-1}) was given from day 2 until killing on day 21. Left column of each pair: refraction data as determined shortly before autopsy; right column: mg Harderian gland per 100 g body weight. * $P < 0.05$; ** $P < 0.01$.

interesting perspectives for pharmacological treatment of juvenile refraction aberrations.

We thank D. Laman for technical assistance. This investigation was supported by a grant from the Netherlands Ophthalmic Research Institute. Ambinon was supplied to us by Organon International BV, The Netherlands.

C. C. KOK-VAN ALPHEN
E. L. NOACH*

Department of Pharmacology,
University of Leiden,
Wassenaarseweg 72,
2333-AL-Leiden, The Netherlands

Received 22 March; accepted 25 October 1977.

*To whom reprint requests should be addressed.

1. Sappey, C. *Gaz. Médic., Paris* 25, 408-411, 423 (1855).
2. Weiss, L. *Anat. Hefte* 8, 193-248 (1897).
3. Scammon, R. E. & Armstrong, E. L. *J. comp. Neurol.* 38, 165-219 (1924/1925).
4. Todd, T. W., Beecher, H., Williams, G. H. & Todd, A. W. *Hum. Biol.* 12, 1-20 (1940).
5. Glickstein, M. & Millodot, M. *Science* 168, 605-606 (1970).
6. Massof, R. W. & Chang, F. W. *Vision Res.* 12, 793-796 (1972).
7. Baylor, E. R. *Nature* 214, 306-307 (1967).
8. Hebel, R. & Stromberg, M. W. in *Anatomy of the Laboratory Rat* 146 (Williams & Wilkins, Baltimore, 1976).
9. Jackson, C. M. *Am. J. Anat.* 15, 1-68 (1913).
10. Hatai, S. *Am. J. Anat.* 15, 87-119 (1913).
11. Donaldson, H. H. & King, H. D. *Am. J. Anat.* 60, 203-229 (1937).

Three forms of delayed skin-test response evoked by mycobacteria

Koch¹ first reported that when guinea-pigs which had been infected 4-6 weeks previously with tuberculosis were challenged intracutaneously (i.c.) with a small number of virulent organisms, there was induration and necrosis within 48 h, and localisation of the challenge dose. This phenomenon (classical tuberculin-type hypersensitivity) has often been regarded as a direct correlate of protective

immunity. But, the same challenge dose was not localised if injected intramuscularly (i.m.). This raised the possibility that 'protection' from the i.c. challenge was a trivial consequence of necrosis and sloughing of the infected skin. The controversy was highlighted by Wilson² who noted that pre-immunised guinea pigs showing an exquisite degree of hypersensitivity at the time of a small i.m. challenge with virulent *Mycobacterium tuberculosis* died more than 40 d before unimmunised controls, whereas weakly hypersensitive animals showed protection. We describe here a form of delayed foot-pad response which is evoked by all the non-pathogenic mycobacteria studied, but differs from the Koch-type response in timing and suppressor cell activity. It is not seen following injection of pathogens, and it is suggested that failure to evoke this response may account for their pathogenicity.

The existence of delayed hypersensitivity responses to mycobacteria, differing from the day 28 'Koch-type' was revealed by skin testing at intervals after injecting 10^6 mycobacteria (see Fig. 1 legend) into groups of 20 or 40 BALB/c mice of either sex. Foot-pads were measured before and 24 h after challenge with $5 \mu\text{g}$ (protein) in $20 \mu\text{l}$ of saline of a $0.22\text{-}\mu\text{m}$ membrane filtered ultrasonicate³ of the same organisms. Delayed hypersensitivity was calculated as per cent swelling in groups of 5-10 mice. All of eight species (provided by Dr J. L. Stanford) known to be non-pathogenic for the mouse (unpublished observation) gave curves of the type shown for *M. nonchromogenicum* in Fig. 1, with a pronounced peak at day 10. The mouse pathogens *M. ulcerans*, *M. kansasii*, and some *M. avium* strains gave a smaller peak between days 5 and 7, but the day 10-11 peak was absent. Similarly, relatively mouse virulent⁴ BCG strains such as Pasteur gave this pathogen-type curve, whereas Glaxo BCG which is of low mouse virulence⁴ gave a powerful day 10 peak. All the organisms gave rise to a later response which resembled the classical Koch type in that the peak occurred between 4 and 6 weeks.

None of the three peaks occurred in T lymphocyte-depleted mice injected with 10^6 *M. nonchromogenicum* or BCG Pasteur. The day 10 response was readily transferred to normal recipients by the intravenous injection 2×10^7 cells from nodes draining the injection site. This transfer was completely abrogated by killing the donor T cells with anti-Thy 1.2, and complement.

Treatment of the mice with intraperitoneal cyclophosphamide, 200 mg per kg body weight 24 h before injection of the bacilli increased the day 6-7 response, without affecting the day 10 peak, if present, or the subsequent day 17 trough. This observation suggested that the day 6-7 peak represented the 'Jones-Mote' phenomenon which typically occurs at 7 d, and is known to be induced by most, if not all mycobacterial species⁵. A regulator of this phenomenon has been reported to be a cyclophosphamide sensitive lymphocyte, in both the guinea pig⁶ and the mouse⁷.

The nature of the 'trough' occurring 17 d after subcutaneous (s.c.) injections of non-pathogenic species (Fig. 1) was investigated by transferring 0.5 ml of fresh serum, or 1.5×10^7 cells from nodes draining the injection site of such day 17 animals, into recipients at other points on the time course (Table 1). Thus normal animals, or mice which had received organisms s.c. 6, 10 or 28 d earlier, received day 17 cells, or serum, or both 1 h before being foot-pad tested.

Neither day 17 cells, nor day 17 serum had demonstrable effects on the foot pad-test responses of control, or day 6 (Jones-Mote) mice. But, day 17 cells suppressed both the day 10, and the day 28 peaks, particularly the latter. Recovery from this suppression was partial by 7 d but complete 14 d later (Table 1).

Day 17 serum enhanced the day 10 peak, but was without

Table 1 Delayed skin-test response to *M. nonchromogenicum*

Treatment of recipients	Controls	Foot-pad responses in recipients (% swelling at 24 h)			Day 28 retested 14 d later
		Day 6	Day 10	Day 28	
None	5.7 ± 4.4	20.3 ± 4.9	38.6 ± 6.2	35.9 ± 12.0	33.1 ± 13.0
1.5 × 10 ⁷ day 17 cells*	4.3 ± 1.6	20.5 ± 4.4	28.3 ± 4.1	9.1 ± 5.4	30.4 ± 7.5
0.5 ml day 17 serum*	5.6 ± 4.4	28.3 ± 7.2	66.9 ± 13.9	35.5 ± 11.2	ND
1.5 × 10 ⁷ cells and 0.5 ml serum*	ND	ND	ND	23.0 ± 8.4	ND

Effect of cells or serum taken from BALB/c mice 17 d after 10⁸ *M. nonchromogenicum* s.c., on the foot-pad responses to 5 µg of antigen, of normal mice, or mice which had received 10⁸ organisms s.c., 6, 10 or 28 d earlier. (10 mice in each group.) ND, not determined.

*Injected intravenously 1 h before the foot-pad test.

†Significantly different from control group. ($P < 0.001$, t -test).

‡Significantly different from control group. ($P < 0.025$).

effect on the day 28 response. But, in subsequent experiments, when transferred into day 28 animals which also received day 17 cells, the serum partially reversed the suppressive effect of the cells, implying that when the serum factor enhanced it did so by opposing suppressor cells, rather than by a direct influence on effector cells. It will be important to find whether this serum factor is similar to that which accompanies tissue-damaging reversal reactions in human leprosy*.

Neither day 17 cells or serum had any significant effect on the delayed hypersensitivity responses to sheep red blood cells, or picryl chloride in appropriately immunised mice.

The major observation here, is that only pathogenic mycobacteria either do not evoke, or are able to suppress the day 10 type of T lymphocyte-dependent response. This is circumstantial evidence for a protective role for this response, which in timing resembles delayed hypersensitivity evoked by other intracellular parasites such as *Listeria monocytogenes*⁹, and *Leishmania*¹⁰. Thus the mechanism of immunity to mycobacteria in the subclinically infected individual may not be the 4–6 week Koch type of hypersensitivity as has often been assumed^{2,11} but may be very similar to that involved in protection from other intracellular organisms. This day 10 peak is accompanied by, and eventually switched off by suppressor

or 'regulator' cells (day 17) the effects of which are modulated by a serum factor. The later day 28 response differs not only in timing, but also in that this type of suppressor cell seems to be absent. This 4–6 week 'Koch-type' of hypersensitivity might be due to a failure of the normal homeostatic suppressor cell mediated regulatory mechanisms. (An ability of mycobacteria to interfere with suppressor cell production would explain the efficiency with which mycobacterial adjuvants will induce auto-immune responses.)

The 4–6 week delay before the appearance of the 'Koch-type' hypersensitivity is explained by the observation that its development is a two stage process. The first stage, which is not antigen specific, requires 2–3 weeks of exposure of lymphoid tissue to mycobacterial components. Thus when sheep red cells (SRBC) were injected into sites previously primed with Pasteur BCG which gives a pathogen-type of curve and fails to evoke day 10 peak, maximal potentiation of the response was seen if the SRBC were injected 2 weeks after the BCG¹². No potentiation was seen if both injections were given at the same time. Similarly, dead mycobacteria, or killed organisms, which traditionally do not induce 'Koch-type' hypersensitivity, give powerful responses if injected twice into the same site with a 2-week interval (data not shown).

Live, virulent organisms do not need to be injected twice, because having escaped the day 10 response by suppressing or failing to evoke it, they continue to proliferate and release antigen which reaches the lymphoid tissue after the first stage of non-antigen specific 'conditioning'. The resulting unregulated, day 28 or 'Koch-type' tissue-destructive hypersensitivity was probably responsible for the accelerated death of Wilsons guinea pigs, and this would help to explain cavitating pulmonary tuberculosis and reversal reactions and nerve damage in tuberculoid leprosy.

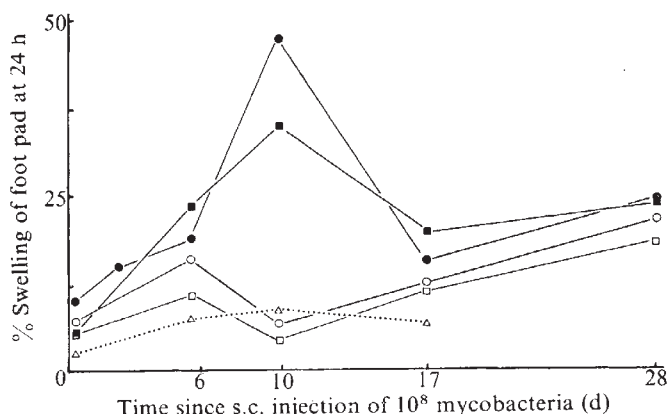
G. A. W. ROOK

School of Pathology,
Middlesex Hospital Medical School,
Ridinghouse Street, London W1, UK

Received 13 July; accepted 31 October 1977.

- Koch, R., *Dtsch. Med. Wschr.* 17, 101–102 (1891).
- Wilson, G. S., Schwabacher, H. & Maier, I. *J. Path. Bact.* 50, 89–109 (1940).
- Stanford, J. L. & Gunthorpe, W. J. *Br. J. exp. Path.* 52, 627–637 (1971).
- Mackness, G. B., Auclair, D. J. & Lagrange, P. H. *J. natn. Cancer Inst.* 51, 1655–1667 (1973).
- Kazda, J. *Zbl. Bakt. Hyg., I. Abt. Orig.* A237, 81–89 (1977).
- Katz, S. I., Parker, D. & Turk, J. L. *Nature* 248, 612–614 (1974).
- Ohmichi, Y., Nomoto, K., Yamada, H. & Takeya, K. *Immunology* 31, 101–110 (1976).
- Biune, G. & Barnetson, R. *St C. Clin. exp. Immun.* 26, 397–402 (1976).
- Mackness, G. B. *J. exp. Med.* 116, 381–406 (1962).
- Preston, P. M., Carter, R. L., Leuchars, E., Davies, A. J. S. & Dumonde, D. C. *Clin. exp. Immunol.* 10, 337–357 (1972).
- Mackness, G. B. *Am. Rev. resp. Dis.* 97, 337–344 (1968).
- Miller, T. E., Mackness, G. B. & Lagrange, P. H. *J. natn. Cancer Inst.* 51, 1669–1676 (1973).

Fig. 1 Mycobacteria were collected from early log phase cultures on solid Sauton's medium, except for Glaxo BCG which was collected by centrifugation from Glaxo liquid production medium. Organisms were dispersed using a 1-ml syringe, washed three times in saline, centrifuged at 600g for 2 min to remove aggregates of organisms, and 5 µl of suitable dilutions spread over a known area of a microscope slide were counted using an auramine stain and a fluorescent microscope. Mice received 10⁸ organisms s.c. in the base of the tail. Delayed hypersensitivity was assessed by the foot-pad test as described in the text. There are 5–10 mice per point. *M. nonchromogenicum* (●), Glaxo BCG (■), Pasteur BCG (□), *M. kansasii* (○), in normal mice, and *M. nonchromogenicum* in adult thymectomised, 750-rad irradiated, bone-marrow reconstituted mice (△).



Influenza A viruses of the H2N2 subtype are lymphocyte mitogens

STUDIES of the effects of viral infections have largely been concerned with intracellular interactions between viral and cellular functions. There is increasing realisation that the cell surface has a major role in cellular function and various cell surface active substances, for example, concanavalin A (con A)¹ and phytohaemagglutinin (PHA)², exert profound effects on cellular metabolism. As the cell surface possesses receptors specific for various viruses and viral antigens, it might be expected that certain viruses could affect cellular function as a result of cell surface interaction. Proliferation is a characteristic response of lymphocytes exposed to these cell surface active compounds. This response can be readily quantitated using ³H-thymidine incorporation. We report here that influenza viruses of the H2N2 subtype are potent mitogens for both B and T lymphocytes³. This finding arose during studies aimed at developing assays for cell-mediated immunity to influenza virus infection⁴⁻⁷. We observed that spleen cells from normal mice showed increased levels of ³H-thymidine incorporation when exposed *in vitro* to the A/Singapore/1/57 virus. This model system should provide insight to questions of lymphocyte triggering and bring attention to this aspect of cell-virus interaction.

As illustrated in Fig. 1, maximum stimulation of normal mouse splenic lymphocytes occurred 48 h after the addition of 10 μ l of A/Singapore virus from a stock containing 10^{7.3} 50% egg infectious dose (EID₅₀) per 0.1 ml. Experiments were designed to determine whether the mitogenic activity of the egg grown stock was virus-associated and, if so, whether it was dependent on virus infectivity. A/Singapore virus was therefore sedimented in an ultracentrifuge and the supernatant and viral pellet were separated. Mitogenic activity was associated with the viral pellet (stimulation ratio (SR) of 3.9) but not with the supernatant fluid (SR of 0.9). In addition, the A/Singapore virus stock was shown to be free from bacterial contamination and had no detectable endotoxin (lipopolysaccharide), a known B cell mitogen, when assayed by the limulus lysate test (sensitivity < 0.5 ng ml⁻¹)⁸. Experiments were performed with live (10^{7.3} EID₅₀ per 0.1 ml) and ultraviolet-inactivated (< 10^{0.5} EID₅₀ per 0.1 ml) virus. These experiments indicated that virus infectivity was not required for stimulation, the SR of live and inactivated virus being 11.6 and 18.3 respectively.

The nature of the cells responding to A/Singapore virus was then evaluated. The response of untreated spleen cells was compared with that of either spleen cells depleted of T cells by treatment with anti- θ serum plus complement or

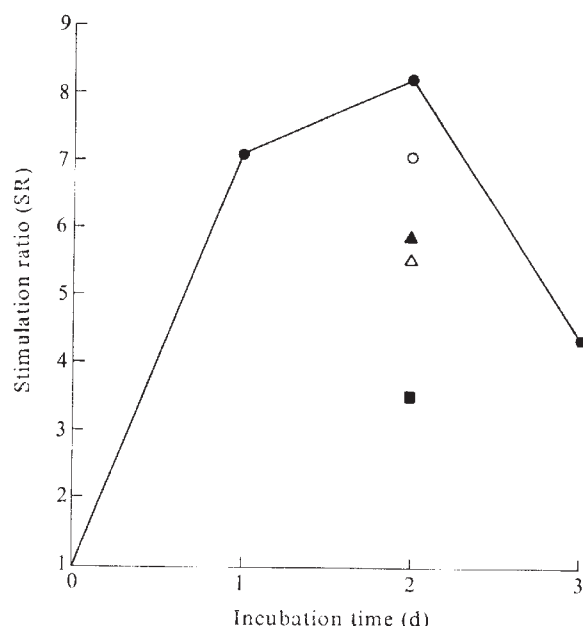


Fig. 1 Effect of time of incubation and dose of A/Singapore virus on *in vitro* proliferation of normal BALB/c splenic lymphocytes. ▲, 0.1 μ l; ●, 10 μ l; △, 30 μ l; and ■, 60 μ l of the A/Singapore virus; ○, 5 μ g con A. In these experiments normal spleens were removed from various animals and single cell suspensions were prepared. Erythrocytes were lysed by exposure of the spleen cells to buffered ammonium chloride solution⁸. Lymphocytes were resuspended to a concentration of 3×10^6 viable cells per ml in RPMI 1640 containing 4% heat-inactivated foetal bovine serum and 2×10^{-4} M 2-mercaptoethanol. Volumes of 0.2 ml of lymphocyte suspension were added to wells of microtitre plates (Microtest II, Falcon). A/Singapore virus was added in varying volumes to quadruplicate samples of normal lymphocytes. These studies were performed using an egg-passaged pool of A/Singapore virus with an egg-infectivity titre of 10^{7.3} EID₅₀ per 0.1 ml. The microtitre plates were incubated for 44 h at 37 °C in a 5% CO₂-95% air atmosphere. At this time 1 μ Ci of ³H-thymidine (specific activity, 6 Ci mmol⁻¹) (Schwarz/Mann) was added per well. After 4 h of further incubation, the cells were collected on to glass fibre filter strips using a MASH II harvester (Microbiological Associates) and the incorporation of ³H-thymidine determined by scintillation counting. Mean incorporation was determined for each group and a stimulation ratio (SR) derived by dividing the mean c.p.m. incorporated in the presence of virus by the mean c.p.m. incorporated in the absence of virus.

spleen cells depleted of B cells by passage through nylon fibre columns according to the method of Julius *et al.*¹⁰. As indicated in Table 1, both T- and B-enriched lymphocyte populations were stimulated by the A/Singapore virus. The responsiveness of B lymphocytes was confirmed in studies using splenic lymphocytes obtained from T-cell deficient nude (athymic) mice (Table 2).

Blastogenic responses were observed with lymphocytes from several mouse strains tested, including those from BALB/c mice raised in a germ-free environment. In addition, lymphocytes from human peripheral blood and spleen cells from Hartley guinea pigs and New Zealand white rabbits were stimulated by A/Singapore virus (Table 2).

Additional influenza viruses were tested for mitogenic activity. Splenic lymphocytes from normal BALB/c mice showed significantly increased levels of ³H-thymidine incorporation when exposed *in vitro* to all influenza viruses of the H2N2 subtype tested (Table 3). Influenza viruses of other subtypes produced little or no stimulation.

These studies indicate that several egg grown influenza viruses of the H2N2 subtype possess significant mitogenic activity for normal T and B lymphocytes. In addition, A/Japan/305 virus grown in a canine kidney cell line (MDCK) possessed mitogenic activity similar to that of the egg-grown virus. The observed blastogenesis is unlikely to be due to a specific immune response associated with cross-

Table 1 Mitogenic effect of A/Singapore influenza virus on T and B enriched lymphocyte populations

Treatment of normal BALB/c spleen cells	Mean stimulation ratio in cultures containing	
	A/Singapore virus	Con A
Untreated	6.2	2.7
Passed through nylon fibre column	4.4	16.7
Treated with complement	2.4	4.1
Treated with anti- θ plus complement	2.4	1.1

Quadruplicate cultures of treated and untreated spleen cells from BALB/c mice were incubated either in the presence or absence of A/Singapore virus as described in Fig. 1. A stimulation ratio for each of the four replicate cultures containing virus was calculated by dividing the counts of ³H-thymidine incorporated in each culture by the mean of the counts of ³H-thymidine incorporated in the quadruplicate cultures not containing virus (approximately 10¹ c.p.m.). The value shown is the mean of the four calculated stimulation ratios. In all experimental groups, the range of the stimulation ratios observed in quadruplicate samples were within 10% of the mean value shown. Similar results were observed in repeat experiments.

Table 2 Mitogenic effect of A/Singapore influenza virus on normal lymphocytes

Source of lymphocytes tested for mitogenic activity	Mean stimulation ratio in the presence of A/Singapore virus
Nude mouse spleen	3.0
C3 H mouse spleen	10.2
Germfree mouse spleen	13.2
Human peripheral blood*	10.6
Guinea pig spleen	4.5
Rabbit spleen	3.2

The stimulation ratios shown are those calculated from four replicate lymphocyte cultures containing A/Singapore virus as described in Table 1. The mean counts of ^3H -thymidine incorporated in quadruplicate cultures not containing virus had a range between 10^3 and 10^4 c.p.m. In all experimental groups, the range of the stimulation ratios observed in quadruplicate samples were within 10% of the mean value shown. Similar results were observed in repeat experiments.

*Healthy adults aged 30–35 yr.

reacting microbial antigens because of the responsiveness of lymphocytes from germ-free mice. Nor is it likely to be due to virus infection of lymphocytes as ultraviolet-inactivated virus retained mitogenicity. Rather the data favour a cellular effect due to virus binding to the lymphoid cell surface. The H2N2 subtype has several distinguishing characteristics which separate those viruses from other influenza subtypes. The most notable feature and basis for classification of the subtype is the possession of the H2 haemagglutinin glycoprotein on the viral envelope. As all A type influenza viruses possess similar matrix and nucleoproteins and because viruses possessing the same neuraminidase (N2) but a different haemagglutinin (H3) did not stimulate lymphocytes (Table 3), these results suggest that the H2 glycoprotein is responsible for this mitogenic activity. Analyses of differences between H2N2 and other influenza viruses should give insight into the molecular basis of the cell surface interaction leading to lymphocyte transformation.

This report presents data that clearly demonstrate that H2N2 influenza viruses are mitogenic for lymphocytes. The possible effects of H2N2 influenza viruses on other types of cells remains to be determined. Another important question is whether other viruses exist with a mitogenic capacity

similar to H2N2 influenza viruses. Finally the biological relevance of virus-stimulated proliferative lymphocyte responses needs to be clarified. Compared to resting lymphocytes, mitogen-activated lymphocytes are more permissive for growth of a wide variety of RNA and DNA viruses^{11–13}. In addition mitogen-induced proliferation has been shown to cause the activation of endogenous xenotropic type-C RNA virus from murine spleen cells^{14–15}. The possible contribution to disease pathogenesis due to mitogenic activity of viruses clearly deserves more study.

G. M. BUTCHKO
R. B. ARMSTRONG
W. J. MARTIN
F. A. ENNIS

Department of Health, Education and Welfare,
Food and Drug Administration,
Bureau of Biologics, Division of Virology,
8800 Rockville Pike,
Bethesda, Maryland 20014

Received 8 September; accepted 1 November 1977.

- Andersson, J., Möller, G. & Sjöberg, O. *Cell Immun.* **4**, 381–393 (1972).
- Doenhoff, M. J., Davies, A. J. S., Leuchars, E. & Wallis, V. *Proc. R. Soc. Lond. B.* **176**, 69–85 (1970).
- Kilbourne, E. D. in *The Influenza Viruses and Influenza* (ed. Kilbourne, E. D.) (Academic, New York, 1975).
- Ennis, F. A., Martin, W. J., Verbonitz, M. W. & Butchko, G. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3006–3010 (1977).
- Butchko, G. M., Martin, W. J. & Ennis, F. A. *Abst. A. Meeting ASM* **317** (1977).
- Ennis, F. A., Martin, W. J. & Verbonitz, M. S. *J. exp. Med.* (in the press).
- Ennis, F. A., Martin, W. J. & Verbonitz, M. W. *Nature* (in the press).
- Roos, D. & Loos, J. A. *Biochim. biophys. Acta* **222**, 565–582 (1970).
- Hochstein, H. D., Elfin, J. F., Cooper, J. F., Seligmann, E. B. & Wolff, S. M. *Bull. Parenteral Drug Ass.* **27**, 139–148 (1973).
- Julius, M., Simpson, E. & Herzenberg, L. A. *Eur. J. Immun.* **3**, 645–648 (1973).
- Bloom, B. R., Himenez, L. & Marcus, P. I. *J. exp. Med.* **132**, 16–30 (1970).
- Wheelock, E. F. & Toy, S. T. *Adv. Immun.* **16**, 123–184 (1973).
- Woodruff, J. F. & Woodruff, J. J. *Prog. Med. Virol.* **19**, 120–160 (1975).
- Greenberger, J. S., Phillips, S. M., Stephenson, J. R. & Aaronson, S. A. *J. Immun.* **115**, 317–320 (1975).
- Moroni, C. & Schumann, G. *Virology* **73**, 17–22 (1976).

Antibody to nuclear ribonucleoprotein penetrates live human mononuclear cells through Fc receptors

It is commonly accepted that antibodies do not penetrate living cells. In only one study anti-purine and anti-nucleoside antibodies were found to penetrate fertilised sea urchin eggs and modify their development¹. Such penetration has been considered unusual and the addition of anti-DNA antibodies does not affect mammalian tissue cells in culture². Direct immunofluorescence of skin biopsies of patients with mixed connective tissue disease (MCTD) using fluorescent anti-IgG has occasionally shown speckled intranuclear fluorescence^{3–5} but it is doubted that IgG entered the cells while still viable. Patients with MCTD have high titres of antibody to nuclear ribonucleoprotein (RNP)^{6–7} which also gives a nuclear speckled pattern on cell substrates in direct immunofluorescence⁸. Should the antibodies to cellular components and nucleic acids which occur in autoimmune diseases be able to penetrate living cells, a novel mechanism of immunologically mediated damage and/or dysfunction could operate. We show here that anti-RNP IgG can penetrate viable human mononuclear cells (MNC), by their surface Fc receptor, and react with their nuclear RNP.

To study the possible penetration of antibodies into viable cells we used an anti-RNP antibody present at high titres (1:16,384,000), as detected by a haemagglutination assay⁷, in the serum of a patient who fulfilled our criteria for MCTD⁹. This serum was unreactive with ribonuclease-treated extractable nuclear antigen (ENA), gave a single precipitin line on Ouchterlony analysis against untreated ENA, indicating its sole reactivity with RNP, and gave the characteristic speckled pattern on indirect immunofluorescence using frozen section of rat stomach as a cell substrate (Fig. 1a). It contained 106,680 mg l⁻¹ of IgG, 2,680 of

Table 3 Mitogenic effect of H2N2 influenza viruses on normal lymphocytes

Viruses tested for mitogenic activity		Mean stimulation ratio
A/Singapore/1/57	(H2N2)*	7.7
A/Japan/305/57	(H2N2)	8.9
A/Japan/170/62	(H2N2)	11.5
A/Taiwan/1/64	(H2N2)	10.7
A/Ann Arbor/7/67	(H2N2)	8.9
A/Tokyo/3/67	(H2N2)	6.4
A/Cornell/1001/67	(H2N2)	9.4
A/New Jersey/8/76	(Hsw1N1)	1.2
A/PR/8/34	(H0N1)	1.0
A/FM/1/47	(H1N1)	1.0
A/Port Chalmers/1/73	(H3N2)	1.1
A/England/42/72	(H3N2)	1.2
A/Victoria/3/75	(H3N2)	0.6

Quadruplicate cultures of spleen cells from BALB/c mice were incubated in the presence of several influenza A virus subtypes as described in Fig. 1. A mean stimulation ratio for each of the four replicate cultures containing virus was calculated as described in Table 1. The mean counts of ^3H -thymidine incorporated in quadruplicate cultures not containing virus was 2×10^4 c.p.m. In all experimental groups, the range of the stimulation ratios observed in quadruplicate samples were within 10% of the mean value shown. Similar results were observed in repeat experiments.

*Major surface antigens: haemagglutinin (H), neuraminidase (N).

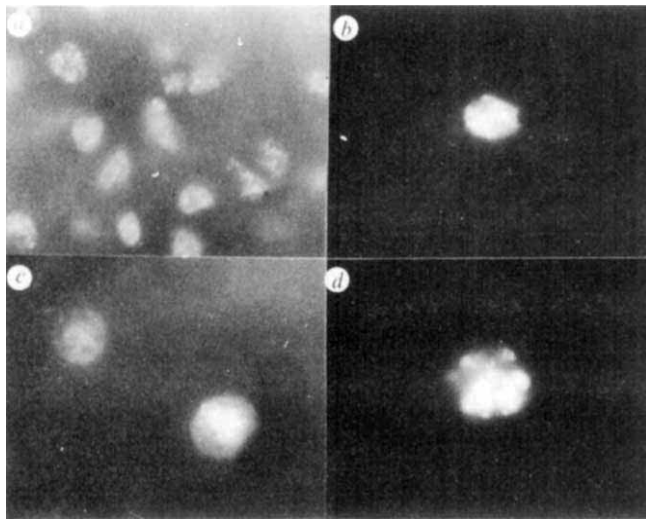


Fig. 1a. Nuclear speckled pattern under epifluorescent microscopy of a frozen section of rat stomach stained with the fluorescein-labelled anti-RNP IgG antibody. The sections were incubated for 25 min with the antiserum, washed for 20 min with isotonic saline under constant agitation, mounted with 0.1 M Glycine buffer, pH 8.6 with 2 parts glycerol, and read. *b*, Speckled nuclear pattern of a viable MNC after incubation with fluorescein-labelled anti-RNP IgG antibody. Viability was controlled by Trypan blue exclusion. *c*, Two viable MNC under fluorescent microscope after incubation with the fluorescent anti-RNP antibody. The lower one shows intranuclear fluorescence as well as three dots of cytoplasmic fluorescence. The upper one shows membrane fluorescence. *d*, A viable MNC, probably a macrophage, shows lumps of intracytoplasmic fluorescence after having been incubated with the fluorescent anti-RNP antibody. All photographs were taken with a C35 47 6070 Zeiss microscope camera using tri-X-pan Kodak film at original magnification of 800 $\times$.

IgA, and 2,520 of IgM as determined by radial immunodiffusion in commercial antibody-agar plates (Behringwerke). The IgG anti-RNP antibody-containing fractions obtained in Sephadex G-200 column fractionation had a titre of anti-RNP antibody of 1:1,048,576 at a protein concentration of 42 mg ml⁻¹. This IgG was tagged to fluorescent isothiocyanate by conventional methods and was absorbed with rat liver powder to eliminate unspecific fluorescence. To avoid the lymphocytotoxic activity found in MCTD sera¹⁰ we also absorbed it repeatedly with human brain cortex, which is known to share antigenic determinants with lymphocytes¹². The brain cortex was previously treated with ribonuclease-A (Sigma, lot 115C-0008) in order to avoid absorption of the anti-RNP antibody by RNP present in this tissue. Following this treatment, the mean lymphocytotoxic activity in a panel from 10 normal human donors in a modification of the microcitotoxicity assay of Terasaki and McClelland^{10,11} was 15%, which is close to that of normal sera¹¹, while the anti-RNP antibody titre remained unchanged. Viability of human peripheral blood mononuclear cells (MNC), obtained from normal donors by the Ficoll-Hypaque method of Böyum¹¹, was tested by Trypan blue exclusion after incubation with the antibody and was found to be 84.6%, which was acceptable for our studies. All subsequent studies were done using this absorbed, fluorescent anti-RNP IgG, and were controlled for viability by Trypan blue exclusion.

One fifth of all viable MNC from five normal subjects showed intranuclear fluorescence (Table 1) with the characteristic speckled pattern of anti-RNP when incubated 60 min at 37 °C with the fluorescent antibody (Fig. 1b). Longer incubation did not increase the number of viable cells with intranuclear speckled fluorescence. Membrane fluorescence (Fig. 1c) was found in 17% of viable cells but seldom concurred with intranuclear fluorescence (Table 1). Capping was seen occasionally. Intracytoplasmic fluorescence seen in 6% of cells (Fig. 1d), was seldom associated with intranuclear fluorescence.

To determine if membrane and cytoplasmic fluorescence could represent steps in the pathway to the nucleus of the anti-RNP antibody, sequential studies were done after increasing in-

cubation times (Table 1). Within 1 h there was a stepwise increase of cells bearing intranuclear fluorescence as well as of those with membrane fluorescence. Inasmuch as fluorescence in these two locations increased in a parallel fashion, it could not be interpreted as steps of the same process. The different proportions of cells bearing nuclear fluorescence, as compared to those bearing membrane fluorescence, when enriched subpopulations were studied (*vide infra*), also suggest that these two different locations of fluorescence occur in different cell subpopulations rather than being stages of antibody penetration. Conversely, cytoplasmic fluorescence decreased along with intranuclear fluorescence in experiments in which penetration of the antibody was blocked (Table 1), suggesting that cytoplasmic fluorescence is due to the antibody on its way to the nucleus.

Five per cent of those viable MNC which adhered to glass wool columns or Petri dishes and stained with myeloperoxidase, indicating them to be predominantly macrophages¹³, were found to have intranuclear staining after incubation with the antibody, while 14% of non-adherent viable MNC showed such intranuclear staining. It therefore seems that the MNC into which the antibody penetrates are primarily lymphocytes, although it may also enter macrophages. To find into which lymphocyte subpopulation the antibody penetrates we prepared cell suspensions enriched for either B or T lymphocytes by a double Ficoll-Hypaque gradient method¹⁴. Suspensions containing predominantly B lymphocytes showed considerably greater percentages of intranuclear staining than those containing predominantly T lymphocytes (Table 1).

The percentages of viable MNC showing intranuclear staining by the labelled anti-RNP antibody are similar to those of MNC bearing receptors for the Fc fragment of IgG¹⁵. To determine if the Fc receptor-bearing MNC are those into which the antibody enters, they were incubated with antibody-coated chicken erythrocytes¹⁶ and the fluorescent anti-RNP antibody simultaneously. Fluorescent staining of nuclei was practically limited to cells forming Fc rosettes.

As the question of viability is crucial in establishing the significance of these findings, and because both Trypan blue exclusion and rosette formation may not be reliable indicators of cell viability at a given moment, we cultured in a 5% CO₂ atmosphere normal MNC pre-incubated with the labelled antibody and shown to have it internalised. Presence of cells bearing speckled intranuclear fluorescence after 24 h of culture confirmed their viability when penetration of the antibody occurred.

Since the antibody penetrates primarily into Fc receptor-bearing cells, this receptor could have a role in the uptake and entrance of the antibody. Normal human MNC were incubated for 1 h with heat-aggregated normal human Cohn's fraction II¹⁷ at concentrations ranging from 16 to 260 μ g per 10⁶ cells before incubation with the labelled antibody (Table 1). Abrogation of antibody penetration by pre-incubation with aggregated γ -globulin was dose related (Table 1), which strongly suggested a blockade of Fc receptors interfering with antibody penetration. Pre-incubation of the MNC cells for 1 h with Fc fragments from human IgG at a concentration of 120 μ g per 10⁶ cells, obtained by papain cleavage of the anti-RNP IgG by a modification of the Porter's method¹⁸, resulted in an 80% decrease of intranuclear fluorescence. Fluorescein labelled F(ab')₂ fragments of the anti-RNP IgG prepared by pepsin digestion¹⁸ failed to penetrate into MNC although they remained capable of giving direct nuclear immunofluorescence in frozen sections of rat stomach.

Because of their large size, antibody coated chicken erythrocytes did not block all Fc receptors and allowed penetration of antibody in experiments combining fluorescent labelling and Fc rosette formation. The dose dependent blocking by aggregated IgG and by Fc fragments supports this explanation.

It is apparent that the anti-RNP IgG antibody penetrates viable human MNC via their Fc receptors and reaches the nucleus where its antigen resides. The significance of these findings may be manifold. (1) The potential for antibodies to penetrate living mammalian cells—this antibody cell penetration could have several implications in immunobiology: selective deletion of the

Table 1 Location of fluorescence in mononuclear cells after incubation with isothiocyanate-labelled anti-RNP IgG antibody in various conditions

Condition	Site of fluorescence, % viable cells		
	Nuclear	Membrane	Cytoplasmic
MNC (whole); preincubation with anti-RNP IgG, time:			
5 min (2)*†	0.5 ± 0.7	1.5 ± 2.1	0
15 min (2)	2.5 ± 0.7	2.0 ± 1.4	2.5 ± 2.1
25 min (2)	5.5 ± 2.1	4.0 ± 1.4	3.0 ± 1.4
40 min (2)	11.5 ± 2.1	16.0 ± 1.4	3.0 ± 1.4
50 min (2)	18.5 ± 2.1	19.0 ± 2.8	3.0 ± 1.4
60 min (5)	19.35 ± 2.8	17.6 ± 6.55	5.8 ± 3.4
120 min (2)	19.5 ± 2.1	17.0 ± 2.8	4.0 ± 2.8
180 min (2)	18.5 ± 2.1	18.0 ± 2.8	4.5 ± 2.1
MNC adherent to glass wool (5)‡§ (83.8 ± 3.8% positive for myeloperoxidase staining)	4.7 ± 1.25	1.2 ± 0.8	1.2 ± 1.2
MNC non-adherent to glass wool (5) (6.6 ± 1.1% positive for myeloperoxidase staining)	13.8 ± 2.5	17.8 ± 2.9	5.3 ± 1.7
B-cell enriched MNC (3)	22.2 ± 4.9	7.7 ± 0.6	9.0 ± 2.6
T-cell enriched MNC (3)	1.4 ± 0.5	13.0 ± 1.4	3.4 ± 1.3
Fc rosette forming MNC (3)	76.5 ± 9.2	0.3 ± 0.6	1.0 ± 1.0
Non-Fc rosette forming MNC (3)	1.0 ± 1.0	12.3 ± 2.1	1.3 ± 1.15
MNC pre-incubated with heat-aggregated Cohn's fraction II (μg per 10 ⁶ cells)			
16 (3)	14.0 ± 3.5	12.3 ± 4.5	9.7 ± 3.2
32 (3)	5.0 ± 2.0	7.0 ± 4.35	5.7 ± 4.0
64 (3)	1.0 ± 1.0	8.3 ± 3.2	0.66 ± 0.6
260 (3)	0.3 ± 0.6	10.7 ± 3.5	0.3 ± 0.6
MNC pre-incubated with Fc fragment of IgG (μg per 10 ⁶ cells)			
60 (2)	10.5 ± 2.1	9.5 ± 0.7	10.0 ± 5.65
120 (2)	4.5 ± 2.1	14.5 ± 6.4	4.5 ± 0.7
MNC incubated with fluorescein labelled F(ab') ₂ (3)	0	6.0 ± 2.6	0

*Parentheses indicate no. of experiments.

†MNC obtained in Ficoll-Hypaque gradient were adjusted to a concentration of 1×10^6 cells per 0.1 ml of minimum essential medium (Eagle) enriched with 0.8% L-glutamine (MEM) were placed in 0.1 ml aliquots and incubated at 37 °C in a 5% CO₂ and 100% humidity atmosphere. Cells were then washed in phosphate buffered saline pH 7.4 (PBS), resuspended in PBS, incubated 5 min with 0.1 ml of a 0.2% Trypan blue dye solution at room temperature, and counted immediately under epifluorescent (Zeiss: Filters UG 1/3, FT. 420, LP 418), and normal illumination. To determine the percentages of cells showing fluorescence at various sites we consider only those cells found to exclude the Trypan blue dye.

‡In this and all subsequent experiments, the incubation time of the MNC with the anti-RNP antibody or its fractions was 60 min.

§Glass wool obtained from J. T. Baker, Phillipsburgh, N. J., USA, Lot 39,284, was loosely packed into a Pasteur pipette filled with MEM and incubated at 37 °C for 30 min. MNC resuspended in the same medium were carefully layered on top of the glass wool and incubated 30 min. with the pipette in vertical position and the tip closed. The pipette tip was opened and 10 ml of MEM were passed through the pipette to collect non-adherent cells. The pipette was then broken and the glass wool washed in PBS, to obtain adherent cells.

||Mean ± 1 σ

antibodies, selective triggering of cell proliferation, a mechanism for memory storing, or, if such be the case, interaction with intracellular antigens. (2) A previously unsuspected role for the Fc receptor in the cellular penetration of antibodies. (3) By a peculiar tropism and/or attraction this antibody, once inside the cell, might make its way to the nucleus to interact with its antigen. And (4), a possible new mechanism of immunologically mediated damage might occur in diseases where antibodies to cell constituents, to nucleic acids and to nucleic acid-protein antigens appear and might modify function when entering a living cell. This mechanism would resemble that of stimulation (type V) as proposed in the Roitt's modification of the Gell and Coombs classification¹⁹.

We have found that intranuclear penetration by antibody occurs *in vivo* in human disease. Patients with MCTD have intranuclear immunoglobulin in their circulating MNC, detectable by means of indirect immunofluorescence using goat antibody to human immunoglobulins, which is thus, also capable of penetrating viable cells. Details of this finding will be reported elsewhere.

We thank the Fondo de Fomento Educativo, Mexico, for financial support.

DONATO ALARCON-SEGIVIA
ALEJANDRO RUIZ-ARGUELLES
EUGENIA FISHBEIN

Department of Immunology and Rheumatology,
Instituto Nacional de la Nutrición,
México 22, D.F., México

Received 9 August; accepted 2 November 1977.

I. Rosenkranz, H. S., Erlanger, B. F., Tanenbaum, S. W. & Beiser, S. M. *Science* **145**, 282-284 (1964).

- Stollar, B. D. in *The Antigens* (ed. Sela, M.) 70 (Academic, New York, 1973).
- Shu, S., Provost, T., Croxdale, M. B., Reichlin, M. & Beutner, E. H. *Clin. exp. Immun.* **27**, 238-244 (1977).
- Gilliam, J. N., Smiley, J. D. & Ziff, M. *Clin. Rev.* **23**, 229A (1975).
- Prystowsky, S. D. & Tuffanelli, D. L. *Proc. XIV int. Congress Rheumatol.* San Francisco, 40 (1977).
- Holman, H. R. *Ann. N.Y. Acad. Sci.* **124**, 800-806 (1965).
- Sharp, G. C., Irwin, W. S., Tan, E. M., Gould, R. G. & Holman, H. R. *Am. J. Med.* **52**, 148-159 (1972).
- Northway, J. D. & Tan, E. M. *Clin. Immun. Immunopath.* **1**, 140-149 (1972).
- Alarcón-Segovia, D. *J. Rheumatol.* **3**, 191-195 (1976).
- Díaz-Jouanen, E., Llorente, L., Ramos-Niembro, F. & Alarcón-Segovia, D. *J. Rheumatol.* **4**, 4-10 (1977).
- Böyum, A. *Scand. J. Lab. clin. Invest.* **21**, Suppl. 97, 77-89 (1968).
- Reif, A. E. & Allen, J. M. V. *J. exp. Med.* **120**, 413-424 (1964).
- Kaplow, L. S. *Blood* **26**, 215-219 (1965).
- Mikulski, S. & Billing, R. *Cell. Immun.* **28**, 69-74 (1977).
- Díaz-Jouanen, E., Rivero, S. J., Llorente, L. & Alarcón-Segovia, D. *Proc. XIV Int. Congress Rheumatol.* San Francisco, 132 (1977).
- Samarut, C., Brochier, J. & Revillar, J. P. *Scand. J. Immun.* **5**, 221-231 (1976).
- WHO/IARC Workshop on human B and T cells, London, 15-17 July 1974 *Scand. J. Immun.* **3**, 521-536 (1974).
- Stanworth, D. R. & Turner, M. W. in *Handbook of Experimental Immunology* (ed. Weir, D. M.) 2nd edn, 10.11-10.20 (Blackwell, Oxford, 1973).
- Roitt, I. *Essential Immunology* 2nd edn, 155-57 (Blackwell, London, 1974).

Carcinogen-induced chromosome breakage in Fanconi's anaemia heterozygous cells

FANCONI'S anaemia (FA) is one of the genetic syndromes in which chromosomal instability has been associated with an increased predisposition to cancer¹. An increased risk of malignancy has also been reported for presumptive carriers of the FA gene^{2,3}, although no evidence of chromosome abnormalities has been found in cultured lymphocytes or fibroblasts from the FA heterozygote. The nature of the genetic defect in FA is still unclear, although there is evidence for defective DNA repair⁴⁻⁶.

Table 1 DEB-induced chromosome breakage in Fanconi's anaemia heterozygous and normal cells

Cell strain	Passage no.	Chromatid breaks	Chromosome breaks	Fragments + deletions	Rearrangements*	No. of breaks per cell†
FA heterozygotes‡						
C13						
Untreated	8	3	0	1	0	0.04
Treated	8	7	7	3	2	0.21
C59						
Untreated	12	0	0	0	0	0
Treated	12	8	1	3	7	0.26
C57						
Untreated	12	6	0	0	0	0.06
Treated	12	10	2	9	2	0.25
Normals‡						
C40						
Untreated	12	2	2	2	1	0.08
Treated	12	5	0	0	1	0.07
C39						
Untreated	11	2	0	1	0	0.03
Treated	11	2	0	3	0	0.05

*Rearrangements include translocations, dicentric, and rings and are scored as two breaks.

†100 cells were analysed in each set.

‡FA heterozygous and normal cultures were not identified until after data were tabulated. C13, 31-yr-old male; father of FA patient C12, previously tested using this same *in vitro* system¹⁰; C59, 40-yr-old male; C57, 49-yr-old female; C40, 23-yr-old male; C39, 26-yr-old male.

Lymphocytes from FA homozygotes have an increased incidence of chromosome aberrations after exposure to ionising radiation⁷ or alkylating agents^{8,9}. We reported recently that baseline breakage rates for FA cell strains varied from 0.20 to 0.36 breaks per cell, and that after carcinogen treatment of the type described below, these rates were increased three to fivefold¹⁰. The same carcinogen treatment had no clastogenic effect on normal, xeroderma pigmentosum and trisomy 18 fibroblasts. There is at present no way to distinguish FA heterozygous individuals from normal. The only distinction that has been reported has been an increased susceptibility to *in vitro* transformation by the oncogenic virus SV40 in both homozygous and heterozygous FA fibroblasts¹¹, but similar results have been obtained with cell strains from patients with several other syndromes. As the ability to identify carriers of the FA gene would be of great benefit because of their increased cancer risk, we report here the results of experiments in which we were able to distinguish FA heterozygous cell strains from normals, after exposing the cells to the difunctional alkylating agent diepoxybutane (DEB), an active mutagen and carcinogen¹²⁻¹⁵.

Early passage skin fibroblasts from five individuals were obtained from Dr M. Swift (University of North Carolina at Chapel Hill). We received the flasks coded, and the identity of the strains was revealed only after results of the experiments were tabulated. Cells from each strain were exposed to 0.01 µg of DEB per ml medium for 6 d, while replicate cultures served as untreated controls. At this concentration, DEB had no cytotoxicity in these strains; treated cells showed approximately the same growth curve as untreated.

The effect of DEB treatment on chromosome breakage in these cells is shown in Table 1. One hundred cells were analysed in each group. The approximately fivefold increase in breakage in the FA heterozygous cells after treatment with DEB was highly significant ($P < 0.001$). The breakage was not concentrated in a few cells, but fit a Poisson distribution. Qualitatively, the chromosomal aberrations induced by DEB in FA heterozygous cells were similar to those which occur spontaneously in FA homozygous cells; open chromatid breaks were the most common finding. In contrast to this result in the FA heterozygous cells, there was no increase in chromosome breakage in normal cells, even those with the highest intrinsic break rate (normal cell strain C40, see Table 1).

Our results indicate that the use of a clastogenic agent in an *in vitro* stress system may have value as a method for determining whether presumptive heterozygotes are in fact gene carriers for FA. In earlier studies exposure of FA heterozygous cells to alkylating agents did not induce increased chromosomal aber-

rations, or increased sister chromatid exchanges greater than that observed in normal cells^{5,8}. Those studies involved direct analysis of exposed lymphocytes. The fibroblast system used in this study may offer a more sensitive means of detecting susceptibility to induced chromosome breakage. Chronic exposure to a non-toxic concentration of the alkylating agent may produce a greater load of DNA damage in a cell population still capable of mitotic division. Perhaps FA heterozygous cells have a reduced capacity to repair DNA, which is not apparent except under the stress of increased damage. As in xeroderma pigmentosum cells^{16,17}, increased sensitivity to chemically-induced chromosome instability occurred in cells lacking spontaneous chromosome breakage.

Although individuals affected with FA are rare, the heterozygote frequency is estimated at 1 in 300 and FA heterozygotes could comprise 1% of all persons dying from cancer². The ability to identify FA heterozygotes would make possible a direct assessment of the risk for carriers of the FA gene acquiring cancer, as compared to the general population. It would also be possible to offer a screening programme for early detection of cancer to individuals identified as FA heterozygotes. The use of an *in vitro* system based on extended exposure to low levels of a carcinogen may provide an effective stress test for the detection of other genes which predispose to cancer.

We thank Drs J. Mitra and R. S. K. Chaganti for helpful discussions, Dr M. Swift for cell strains, and Dr B. M. Goldschmidt for providing the DEB. This study was supported in part by the Martha and Regina Brand Foundation.

ARLEEN D. AUERBACH

Laboratory of Genetics,
Memorial Sloan-Kettering Cancer Center,
New York, New York 10021

SANDRA R. WOLMAN

Department of Pathology,
New York University School of Medicine,
New York, New York 10003

Received 29 August; accepted 1 November 1977.

- German, J. *Prog. med. Genet.* **8**, 61-101 (1972).
- Swift, M. *Nature* **230**, 370-373 (1971).
- Swift, M., Cohen, J. & Pinkham, R. *Am. J. hum. Genet.* **26**, 304-317 (1974).
- Poon, P. K., O'Brien, R. L. & Parker, J. W. *Nature* **250**, 223-225 (1974).
- Latt, S. A., Stetten, G., Juergens, L. A., Buchanan, G. R. & Gerald, P. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 4066-4070 (1975).
- Rensen, J. F. & Cerutti, P. A. *Proc. natn. Acad. Sci. U.S.A.* **73**, 2419-2423 (1976).
- Higashimura, M. & Conen, P. E. *Blood* **38**, 336-342 (1971).
- Schuler, D., Kiss, A. & Fabian, F. *Humangenetik* **7**, 314-332 (1969).
- Sasaki, M. S. & Tonomura, A. *Cancer Res.* **33**, 1829-1836 (1973).

10. Auerbach, A. D. & Wolman, S. R. *Nature* **261**, 494–496 (1976).
11. Todaro, G. J., Green, H. & Swift, M. R. *Science* **153**, 1252–1254 (1966).
12. Kolmark, G. & Westergaard, M. *Hereditas* **39**, 209–224 (1953).
13. Bird, M. J. & Fahmy, O. G. *Proc. R. Soc. B* **140**, 556–578 (1953).
14. Van Duuren, B. L. *Ann. N. Y. Acad. Sci.* **163**, 633–651 (1969).
15. Wolman, S. R. & Sivak, A. *Lab. Invest.* **33**, 670–677 (1975).
16. Sasaki, M. S. *Mutat. Rev.* **20**, 291–293 (1973).
17. Stich, H. F., Stich, W. & San, R. H. C. *Proc. Soc. exp. Biol. Med.* **142**, 1141–1144 (1973).

Increased CTP synthetase activity in cancer cells

AN insight into the biochemical strategy of cancer cells has been achieved by the conceptual and experimental approach provided by the molecular correlation concept^{1,2}. Studies carried out using this approach with model systems of rat hepatomas and kidney tumours of different growth rates revealed a biochemical imbalance in cancer cells, which manifested itself in progressive changes in activities of key enzymes and overall metabolic pathways that correlated with tumour growth rate^{1,2}. Some of these biochemical alterations have been shown to apply to human primary liver³ and kidney⁴ carcinomas. Those alterations in gene expression that are linked to the increase in the expression of malignancy are manifested in the increased activities of key glycolytic-, purine-, pyrimidine-, DNA- and polyamine-synthesising enzymes^{1–8}. Concurrently, decreases occur in the activities of the key enzymes of gluconeogenesis, purine and pyrimidine catabolism and of the urea cycle^{1–5,9,10}. In addition to such a progression-linked (growth-rate-linked) metabolic imbalance, the reprogramming of gene expression in cancer cells entails transformation-linked alterations that are present in all hepatomas irrespective of growth rate or extent of differentiation^{1,2,11–15}. Here we report that CTP synthetase (UTP:L-glutamine amido ligase, EC 6.3.4.2) increased in all the hepatomas examined, the activity being highest in the rapidly growing tumours. Thus, in liver neoplasia the activity of this enzyme is both transformation- and progression-linked. CTP synthetase activity was also markedly increased in transplantable kidney tumours in the rat and in primary renal cell carcinomas in man.

Our studies on CTP synthetase, the enzyme that catalyses the synthesis of CTP from UTP, were undertaken because the activity of UDP kinase, an enzyme that produces UTP, was elevated in all hepatomas examined¹⁵. In rapidly growing hepatoma 3924A, however, the concentration of UTP was unaltered, but the CTP concentration was increased fivefold¹⁶. An accumulation of CTP might occur because of decreased utilisation, which is not likely in the rapidly growing tumour; or it might be caused by an increase in the activity of CTP synthetase, an enzyme that is a sensitive target for anti-tumour active glutamine antagonists¹⁷. We have tested the hypothesis that the observed increase in CTP concentration may be accounted for, at least in part, by an increase in CTP synthetase activity. Our systematic studies revealed that CTP synthetase is the rate-limiting enzyme in hepatic CTP biosynthesis (Fig. 1). The specific activity of CTP synthetase was increased in all hepatomas examined, and the elevation correlated with the increase in tumour growth rate. This increased potential for CTP biosynthesis provides one mechanism by which the cancer cell can adjust the availability of CTP for RNA synthesis and for the formation of 2'-deoxy CTP for the synthesis of DNA.

Male ACI/N and Buffalo inbred strains of rats and Wistar rats were maintained as described previously³. The procedures for obtaining tumours and regenerating livers have been described³. The growth rates of the various hepatoma lines ranged from 2–52 weeks, as measured by the amount of time required for the neoplasms to reach a diameter of 1.5 cm.

The kinetic characteristics of CTP synthetase in crude enzyme preparations and in partially (50-fold) purified extracts of rat liver and rapidly growing hepatoma 3924A were similar. The pH optima were broad, ranging from 7.0–9.0. The apparent K_m values for the ligands UTP, ATP, glutamine and GTP at pH 7.4 and 37 °C were 0.3, 0.6, 0.1, and 0.07 mM, respectively, both in liver and hepatoma 3924A crude extracts.

The specific activity of CTP synthetase in the liver of freely fed rats ranged from 4–6 nmol per h per mg protein. A comparison of

the activities of enzymes involved in the *de novo* synthesis, salvage or recycling of UMP, and in the conversion of UMP to CTP (Fig. 1), shows that CTP synthetase activity is the lowest^{15,18,19}.

In the liver tumours of slow and medium growth rate the specific activity of CTP synthetase was increased 1.8- to 3.6-fold that of normal rat livers, whereas the activity was increased from 4.6- to 11.2-fold in rapidly growing hepatomas (Table 1).

In 24-h regenerating livers, CTP synthetase activity increased twofold that observed in the liver of sham-operated control rats. This agrees with observations indicating an increase in uridylate biosynthesis in regenerating liver²⁰. CTP synthetase activity in the average liver cell of the 5-d-old rat was 31% of that of normal adult rat. Since increased CTP synthetase activity was present in the hepatomas, but not in the rapidly growing differentiating liver, CTP synthetase activity is not linked with differentiation. Since the regenerating liver has a proliferation rate similar to that of rapidly growing hepatoma 3924A (ref. 1), the elevated activity in this hepatoma (sevenfold) was higher than that in regenerating liver (twofold).

Studies in this laboratory demonstrate a marked step-up in hepatoma CTP biosynthetic capacity^{15,18}. The key enzymes participating in the *de novo* biosynthesis of UMP were increased in parallel with the increase in CTP synthetase activity and with tumour growth rate¹⁸. The salvage enzymes, uridine phosphorylase, uridine kinase, and uracil phosphoribosyltransferase, were similarly elevated¹⁹. UDP kinase was increased in all tumours irrespective of growth rate¹⁵. In contrast, the activity of the rate-limiting catabolic enzyme, dihydrouracil dehydrogenase, was decreased⁹. Thus, the programme displayed in the cancer cells provides for an increased capacity for CTP biosynthesis and a decreased capability for the catabolism of the precursors. These alterations should confer selective advantages on the cancer cells.

Note that CTP synthetase activity was also increased in slowly growing rat kidney tumour MK-3 (3.5-fold) and in three cases of human renal cell carcinoma (2.5-fold).

Our demonstration of high CTP synthetase activity in hepatomas suggests that the reprogramming of gene expression in malignant transformation and progression is linked with the

Fig. 1 Metabolic pathways for CTP biosynthesis and utilisation. Numbers in parentheses are specific activities of the enzymes (nmol per h per mg protein) at 37 °C and pH 7.4. Further details of the metabolic reactions given in schematic form in this figure are cited in the text^{1,2}. ASP, Aspartate; CP, carbamoylphosphate; OMP, orotidine-5'-monophosphate; UMP, uridine-5'-monophosphate; PRPP, 5-phosphoribose-1-pyrophosphate; U, uracil; UR, uridine; β -Ala, β -alanine; UDP, uridine-5'-diphosphate; UTP, uridine-5'-triphosphate; CTP, cytidine-5'-triphosphate; ATP, adenosine-5'-triphosphate; GTP, guanosine-5'-triphosphate; dGTP, deoxyguanosine-5'-triphosphate; dATP, deoxyadenosine-5'-triphosphate; RNA, ribonucleic acid; DNA, deoxyribonucleic acid.

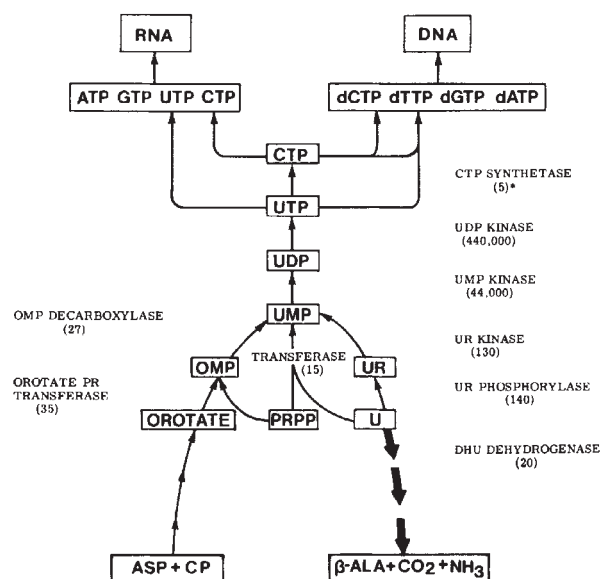


Table 1 CTP synthetase activity in hepatomas of different growth rates

Tissues	Growth rate (months)	Protein		Activity		
		mg g ⁻¹	% Control	nmol per h per mg protein	% Control	
Normal livers from Buffalo rats						
Control for	9618A	100.0±0.6		5.8±0.3		
	28A	111.7±2.0		5.9±0.4		
	8999	100.0±0.6		5.8±0.3		
	7787	111.7±2.0		5.9±0.4		
	44	111.7±2.0		5.9±0.4		
	9633	100.0±0.6		5.8±0.3		
	7777	94.0±1.2		5.1±0.7		
	7288C	98.9±0.9		4.1±0.1		
	9618A2	98.9±0.9		4.1±0.1		
Normal livers from ACI/N rats						
Control for	3924A	95.5±1.2		5.5±0.3		
	3683F	98.4±1.3		5.4±0.2		
Hepatomas						
	9618A	12.4	93.0±4.0	93	19.7±1.4	340*
	28A	6.8	86.9±0.9	78*	21.2±0.6	359*
	8999	6.3	82.2±0.4	82*	10.7±0.4	184*
	7787	6.0	91.7±1.7	82*	11.1±0.5	188*
	44	5.4	98.5±1.5	88*	13.1±0.4	222*
	9633	5.2	76.8±0.8	77*	16.1±0.4	277*
	7777	1.0	59.7±0.7	63*	23.7±1.5	464*
	3924A	0.9	71.0±0.9	74*	38.6±1.4	702*
	7288C	0.8	78.8±0.6	80*	40.7±1.0	999*
	9618A2	0.8	71.7±1.0	72*	37.1±0.8	911*
	3683F	0.5	73.1±0.9	74*	60.6±2.9	1122*

*Significantly different from values of corresponding control liver ($P = < 0.05$).
Data are mean ± s.e. of four or more animals in each group.

Tissue homogenates (15%, w/v) were prepared at 4 °C in 0.15 M KCl, pH 7.4. CTP synthetase activity was determined in the 100,000g supernatant fluids by measuring the conversion of 4-¹⁴C-UTP to 4-¹⁴C-CTP. The components (mM) of the incubation mixture (pH 7.4, 37 °C) were contained in 200 µl: glycylglycine, 70; MgCl₂, 18; beta-mercaptoethanol, L-glutamine and NaF, 10; phosphoenolpyruvate and ATP, 8; GTP, 1; UTP, 2; 4-¹⁴C-UTP (60 mCi mmol⁻¹), 0.03; and supernatant fluid, 100 µl. The assay was initiated by addition of supernatant fluid and terminated at 5-min intervals by adding perchloric acid and immersing the reaction tube in an ice-bath. Samples were neutralised with ethylenediamine tetraacetic acid-KOH solution and the potassium perchlorate removed by centrifugation. Ten microlitres of the clear supernatant fluid was applied on PEI-cellulose plates (Brinkmann) together with 20 nmols of a carrier mixture containing the mono-, di- and triphosphates of uridine and cytidine. Ascending chromatography was carried out at 4 °C in rectangular glass tanks which contained 100 ml 0.8 M ammonium sulphate as described previously²¹. CTP was located approximately 2 cm below the UTP spot as determined using ultraviolet light and autoradiography. Ultraviolet absorbing spots were cut out and placed into counting vials with scintillation fluid (Econofluor, New England Nuclear). Radioactivity was determined in a Packard Tri-Carb (Model 3390) liquid scintillation spectrometer. Protein concentrations were assayed²², using recrystallised bovine serum albumin as standard. CTP synthetase activity was linear with amount of enzyme added and time elapsed in the assay; specific activity was expressed as nmol substrate metabolised per h per mg protein.

expression of the CTP-synthesising enzyme activity. In this laboratory, 12 such transformation-linked increases in enzyme activities have so far been discovered, and they all relate to an increased potential in the channelling of metabolites to strategic biosynthetic processes. The increase in the activities of glucose-6-phosphate dehydrogenase (EC 1.1.1.49) and transaldolase (EC 2.2.1.2) provides an increased capacity for channelling glycolytic intermediates into pentose phosphate biosynthesis. Elevations in the activities of phosphorylribose-1-pyrophosphate (PRPP) synthetase (EC 2.7.6.1)¹² and glutamine PRPP amidotransferase (EC 2.4.2.14)¹³ provide an increased potential for utilising ribose-5-phosphate and PRPP for *de novo* purine biosynthesis. Increases in the activities of adenylosuccinate synthetase and adenylosuccinase should yield an increased capacity for the *de novo* production of adenine nucleotides³. The elevated activity of AMP deaminase (EC 3.5.4.6)³ and IMP dehydrogenase (EC 1.2.1.14)⁶ provide an increased capability for IMP formation and its utilisation in the *de novo* biosynthesis of guanine nucleotides. The increased activities of uridine kinase (EC 2.7.1.48), uridine phosphorylase (EC 2.4.2.3), uracil phosphoribosyltransferase (EC 2.4.2.9)¹⁹ and UDP kinase¹⁵, provide an increased capacity for UTP biosynthesis. In contrast, the activities of certain catabolic enzymes are decreased in the hepatomas, including inosine phosphorylase (EC 2.4.2.1)¹⁹, xanthine oxidase (EC 1.2.3.2), uricase (EC 1.7.3.3) and thymidine phosphorylase (EC 2.4.2.4)¹⁴. The display of this transformation-linked programme indicates that an enzymatic imbalance emerged in neoplasia reflecting ordered and reciprocal alterations in gene expression. Some of these enzymes constitute promising points for attack in

the design of selective enzyme-pattern-targeted chemotherapeutic agents and offer opportunities for selectively preventing host toxicity by means of metabolite rescue²³.

This work was supported by the USPHS.

JIM C. WILLIAMS
HARUTOSHI KIZAKI
GEORGE WEBER

Laboratory for Experimental Oncology,
Indiana University School of Medicine,
Indianapolis, Indiana 46202

HAROLD P. MORRIS

Department of Biochemistry,
Howard University College of Medicine,
Washington, DC 20001

Received 15 July; accepted 3 November 1977.

1. Weber, G. *N. Engl. J. Med.* **296**, 486-493 (1977).
2. Weber, G. *N. Engl. J. Med.* **296**, 541-551 (1977).
3. Weber, G., Prajda, N. & Jackson, R. C. *Adv. Enz. Regul.* **14**, 3-24 (1976).
4. Weber, G. *et al.* in *Biological Characterisation of Human Tumours* (eds Davis, W. & Maltoni, C.) 60-72 (Excerpta Medica, Amsterdam, 1976).
5. Weber, G., Goulding, F. J., Jackson, R. C. & Ehle, J. N. in *Biological Characterisation of Human Tumours* (eds Davis, W. & Maltoni, C.) (Excerpta Medica, Amsterdam, in the press).
6. Jackson, R. C., Weber, G. & Morris, H. P. *Nature* **256**, 331-333 (1975).
7. Elford, H. L., Freese, M., Passamani, E. & Morris, H. P. *J. biol. Chem.* **245**, 5228-5233 (1970).
8. Williams-Ashman, H. G., Coppoc, G. L. & Weber, G. *Cancer Res.* **32**, 1924-1932 (1972).
9. Queener, S. F., Morris, H. P. & Weber, G. *Cancer Res.* **31**, 1004-1009 (1971).
10. Weber, G., Queener, S. F. & Morris, H. P. *Cancer Res.* **32**, 1933-1940 (1972).
11. Weber, G., Jackson, R. C., Williams, J. C., Goulding, F. J. & Eberts, T. J. *Adv. Enz. Regul.* **15**, 53-77 (1977).
12. Heinrich, P. C., Morris, H. P. & Weber, G. *Cancer Res.* **36**, 3189-3197 (1976).
13. Prajda, N., Katunuma, N., Morris, H. P. & Weber, G. *Cancer Res.* **35**, 3061-3068 (1975).
14. Prajda, N., Morris, H. P. & Weber, G. *Cancer Res.* **36**, 4639-4646 (1976).
15. Williams, J. C., Morris, H. P. & Weber, G. *Nature* **253**, 567-569 (1975).
16. Jackson, R. C., Boritzki, T. J., Morris, H. P. & Weber, G. *J. Life Sci.* **19**, 1531-1536 (1976).
17. McPartland, R. P., Wang, M. C., Bloch, A. & Weinfeld, H. *Cancer Res.* **34**, 3107-3111 (1974).

18. Sweeney, M. J., Parton, J. W. & Hoffman, D. H. *Adv. Enz. Regul.* **12**, 385-396 (1974).
 19. Weber, G. *et al. Adv. Enz. Regul.* **16**, (in the press).
 20. Bucher, N. L. R. & Malt, R. A. *Regeneration of Liver and Kidney* (Little Brown, Boston, 1971).
 21. Neuhard, J., Randerath, E. & Randerath, K. *Analyt. Biochem.* **13**, 211-222 (1965).
 22. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).
 23. Jackson, R. C., Williams, J. C. & Weber, G. *Cancer Treatment Rep.* **60**, 835-843 (1976).

Central nervous system effects of limb amputation in man

WE report here the results of an electrophysiological study of the central effects of long-standing limb amputation in man. The findings are felt to have significance for an understanding of neurotrophic mechanisms. The results show that, after amputation, the impulse-conducting ability of a peripheral nerve is severely reduced and that this is accompanied by significant changes in the evoked cortical response; in contrast, function in certain reflex pathways is apparently well maintained. This type of study complements a number of early anatomical investigations^{1,2}. In reviewing the early studies it is evident that much disagreement exists concerning the extent of the atrophy to be found in the cells of the anterior and posterior horns, and in the myelinated fibres of the anterior and posterior roots and of the ascending and descending tracts in the spinal cord. There is the added problem, common to all neuropathological studies, of deciding to what extent impulse conduction and synaptic transmission might have been interfered with. By using a functional approach our study has resolved some of these uncertainties.

The study was performed on 10 male patients, aged 11 to 47 yr. Nine patients had suffered amputation of a hand and part of the forearm and in one case a similar loss was the result of a congenital anomaly; the durations of the lesions varied from 4 months to 30 yr. In each patient the central stump of the ulnar nerve was stimulated maximally below the elbow and the compound action potential was recorded between the elbow and axilla. Recordings were also made from the somatosensory 'hand' area of the contralateral cortex using conventional techniques^{3,4}. Finally, the reflex effects of such stimulation were studied on volitional activity in the ipsilateral triceps muscle using rectification and averaging⁵. All the stimulating and recording procedures were performed with surface electrodes; these comprised chlorided silver disks or cups coated with a conducting cream. Control observations were made, using similar electrode placements, on the intact arms of the patients and on both arms (and contralateral somatosensory cortices) of 15 healthy controls. All the patients and control subjects had given informed consent for the procedures.

The results in the amputees were consistently abnormal in several respects (Fig. 1 and Table 1). First, the compound action potential in the ulnar nerve was markedly reduced on the side of the amputation, the mean peak-to-peak amplitude being only 13% of that on the control side. Maximum impulse velocities in the conducting fibres were usually normal, though low velocities in two subjects caused the mean value to be

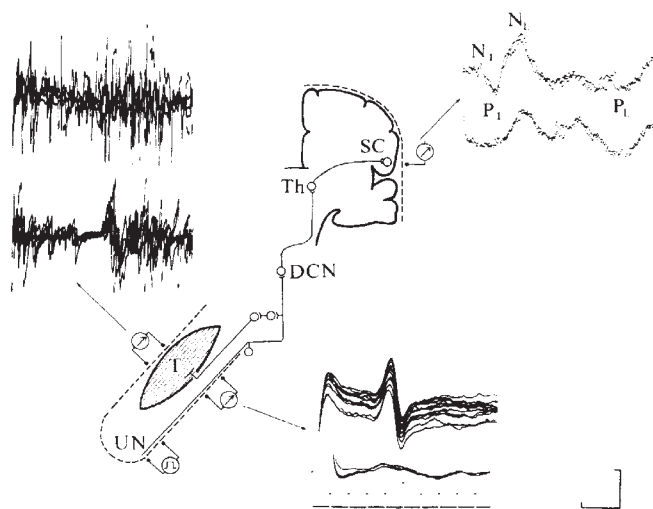


Fig. 1 Experimental arrangement, neural pathways, and results of study in a 32-yr-old man who had suffered traumatic amputation of his right hand 3 yr previously. In each pair of recordings the upper and lower traces show the results of stimulating ulnar nerves (UN) in the intact and partially-amputated arms respectively. Stimuli were delivered at the start of each trace. Recordings from somatosensory cortex (SC) were averaged 32 times in each of two trials. The recordings from the triceps muscle (T) were made during weak voluntary contraction and several traces have been superimposed (rectified and averaged responses not shown). Vertical bar represents 10 μ V, 20 μ V and 40 μ V for recordings from cortex, peripheral nerve and triceps respectively; horizontal bar represents 40 ms except for peripheral nerve traces in which calibration marks denote milliseconds. DCN, dorsal column nuclei; Th, thalamus; triceps reflex pathway simplified.

significantly smaller. Second, the earliest negative (N_1) and positive (P_1) components of the evoked cortical responses were significantly reduced, being undetectable in four patients; in contrast, later components (N_{1-2} , P_{1-2} in Table 1) were better preserved. Finally, volitional activity in the triceps muscle was modulated to a greater extent by inputs from the amputated ulnar nerve stump than from the nerve on the intact side; this applied both to the initial inhibitory phase (40-100 ms post-stimulus) and to the succeeding excitatory phase. Control observations in normal subjects suggested that the reason for the discrepancy was that inhibition had decreased on the intact side of the amputees, presumably as a result of the extra use made of that limb after injury to the opposite arm.

These results suggest that, in normal circumstances, the functional integrity of motor and sensory axons in human limb nerves is heavily dependent on intact connections with muscle fibres and receptor tissues. Despite the severe loss of excitable axons which follows amputation, sufficient sensory nerve fibres can remain functional to elicit normal polysynaptic reflexes (involving ipsilateral motoneurons with intact axons). Finally, it is probable that there is some loss of function in the

Table 1 Comparison of responses evoked in ulnar nerve, ipsilateral triceps muscle, and contralateral somatosensory cortex following maximal stimulation of ulnar nerves in partially-amputated and intact arms of patients

		Control limbs	Patient limbs		P
		(30)	Intact (10)	Amputated (10)	
Nerve	Action potential (μ V)	71.1 $\pm$ 26.8	66.1 $\pm$ 34.1	8.7 $\pm$ 8.6	<0.001
	Conduction velocity (m s ⁻¹)	76.9 $\pm$ 7.6	73.6 $\pm$ 8.6	68.2 $\pm$ 9.6	<0.02
Cortex	N_1 - P_1 response (μ V)	3.9 $\pm$ 1.8	4.3 $\pm$ 1.2	1.4 $\pm$ 1.4	<0.001
	N_{1-2} - P_{1-2} response (μ V)	8.8 $\pm$ 2.8	9.4 $\pm$ 3.8	6.2 $\pm$ 2.5	<0.01
Triceps	Inhibition (%)	75.4 $\pm$ 17.0	41.8 $\pm$ 25.7	79.7 $\pm$ 25.9	>0.5
	Excitation (%)	46.6 $\pm$ 26.0	42.8 $\pm$ 20.9	68.4 $\pm$ 51.1	>0.2

Results in 15 normal subjects are shown for comparison. Parentheses indicate numbers of limbs studied. Significance levels calculated by Student's *t*-test and refer to differences between results of nerve stimulation in partially-amputated limbs and in arms of control subjects.

dorsal column–medial lemniscus pathway subserving the early components of the evoked cortical response, beyond that due to lesions of the peripheral sensory nerve fibres. Further studies, aimed at documenting the time course of the central effects or at preventing their occurrence, would be of considerable interest.

ALAN MCCOMAS
ROBERTO SICA
SIKHAR BANERJEE

MRC Group in Developmental Neurobiology,
Room 4U7, McMaster University Medical Centre,
1200 Main St West, Hamilton, Ontario, Canada

Received 19 August; accepted 4 November 1977.

1. Barratt, J. O. W. *Brain* **24**, 310–328 (1901).
2. Ranson, S. W. *J. comp. Neurol.* **16**, 265–293 (1906).
3. Giblin, D. R. *Ann. N.Y. Acad. Sci.* **112**, 93–142 (1964).
4. Halliday, A. M. & Wakefield, G. S. *J. Neurol. Neurosurg. Psychiat.* **26**, 211–219 (1963).
5. Caccia, M. R. et al. *J. Neurol. Neurosurg. Psychiat.* **36**, 960–977 (1973).

Acetylcholine-induced channels and transmitter release at human endplates

MYASTHENIA gravis, a disease affecting the human neuromuscular junction, is thought to result from a postsynaptic defect, namely a reduction in the number of functional acetylcholine (ACh) receptors (reviewed in ref. 1). Thus at myasthenic endplates the spontaneous miniature endplate potentials (m.e.p.ps) and impulse-evoked endplate potentials (e.p.ps) are reduced in amplitude². Furthermore, the diminished number of α -bungarotoxin binding sites^{3,4} indicates a reduction in the number of ACh receptors in the postsynaptic membrane which leads to a reduced sensitivity to ACh⁵; and preliminary experiments with voltage noise (B. Katz, R. M. & J. Newsom-Davis, unpublished) suggested that the decrease in sensitivity does not involve changes in the characteristics of the channels. We have further examined the properties of channels opened by ACh at normal and myasthenic endplates as well as some characteristics of transmitter release in the two conditions. The experiments described here indicate that the single channel conductance and the mean channel life-time are similar in normal and myasthenic muscle membranes but that there are differences in the Ca-dependence of transmitter release from normal and myasthenic nerve terminals.

Intercostal muscle from myasthenic and control patients were dissected to give a small bundle of muscle fibres, sometimes with its nerve supply. Stimulation of the nerve was via a glass capillary suction electrode. Preparations were perfused with an oxygenated medium and examined using conventional electrophysiological techniques for intracellular recording and voltage clamping. For noise experiments, ACh was applied iontophoretically from a micropipette (25–50 M Ω) positioned between the recording and voltage-clamping microelectrodes.

To study the characteristics of channels induced by ACh we have investigated miniature endplate currents, which result from the action of single packets of transmitter on the muscle membrane, and also the membrane current noise which occurs during the application of ACh⁶. As at the nerve–muscle junctions of other species⁷, m.e.p.ps at voltage-clamped human endplates vary greatly in amplitude and shape, with some m.e.p.ps being clearly composite. The decay of the faster m.e.p.ps was exponential^{8–10}, however, and these were used for subsequent analysis.

Examples of a normal and a myasthenic miniature endplate current are illustrated in Fig. 1 ($V_m = -80$ mV, $T = 23^\circ\text{C}$) which shows a semi-logarithmic plot of their decay as a function of time. Comparison of the time constant of decay of m.e.p.ps ($\tau_{m.e.p.c.}$) from normal and myasthenic endplates reveals that $\tau_{m.e.p.c.}$ is very similar in the two situations; $\tau_{m.e.p.c.}$ (normal) = 2.0 ± 0.1 ms (mean $\pm$ s.e.), $\tau_{m.e.p.c.}$ (myasthenic) = 2.1 ± 0.1 ms. At both normal and myasthenic endplates the duration of m.e.p.ps was prolonged by hyperpolarisation of the membrane, as occurs in frog and toad muscle^{8–10}.

The correspondence of the values for the duration of $\tau_{m.e.p.c.}$ at myasthenic and normal endplates may be taken to reflect a

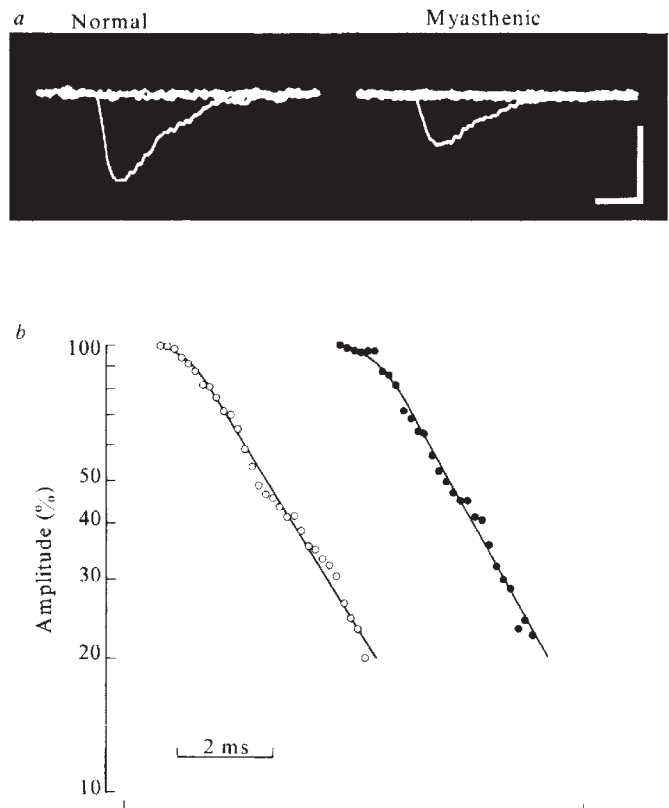


Fig. 1 *a*, Typical miniature endplate currents from normal and myasthenic human endplates. The figures are of several superimposed traces replayed from analog tape. M.e.p.ps were recorded under voltage clamp ($V_m = -80$ mV, $T = 23^\circ\text{C}$) through a 500-Hz low-pass filter which prolongs the rising phase and slightly reduces the amplitude but does not greatly affect the time constant of decay. Calibration 2 ms and 2 nA. *b*, Decay phases of the m.e.p.ps shown in *a*, plotted on semi-logarithmic co-ordinates. Straight lines through the points were fitted by the least squares method to values between 80–20% of the m.e.p.c. amplitude to give the time constant of decay. $\circ$, Normal $\tau_{m.e.p.c.} = 1.9$ ms; $\bullet$, myasthenic $\tau_{m.e.p.c.} = 1.8$ ms.

similarity in the duration of the mean open state of the acetylcholine channel in the myasthenic and normal muscle membrane, providing other factors, such as time for diffusion of acetylcholine from the cleft¹¹, are equal. It has been possible to derive certain of the properties of the single endplate channel by analysing the voltage-clamped noise which occurred during the steady iontophoretic application of ACh to endplates. In similar experimental conditions, iontophoretic application of ACh produced consistently smaller mean currents at myasthenic than at normal endplates, indicating a reduced postsynaptic sensitivity to ACh.

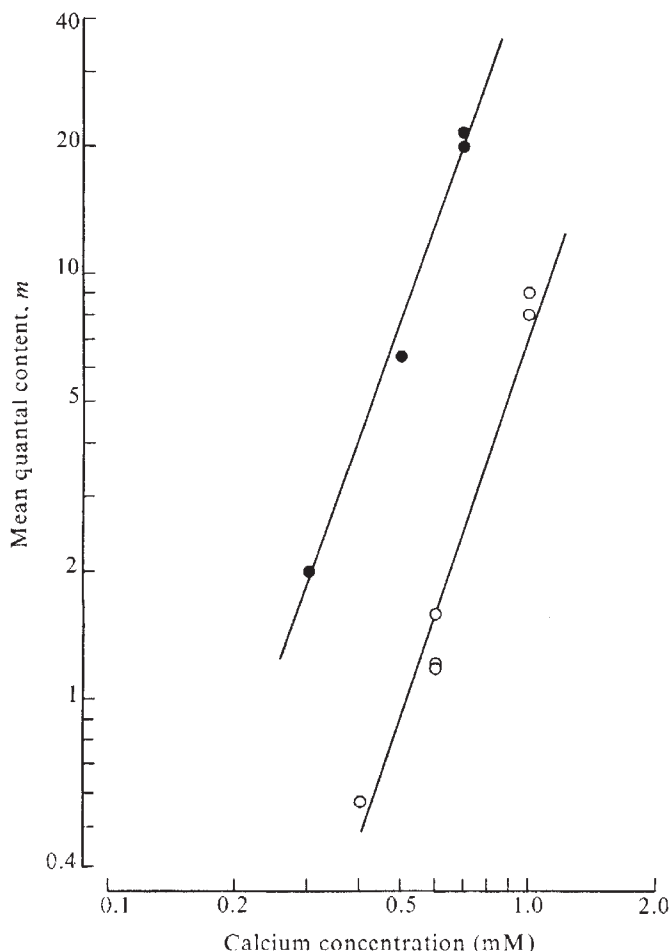
The mean life-time of the channel (τ_{noise}) was obtained^{6,12} from the cutoff frequency (f_c) of the power spectrum (the frequency at which the spectrum has decreased to half of the zero frequency asymptote) by fitting a single Lorentzian component of the form $S(f) = S(0)/(1 + (f/f_c)^2)$ to the spectrum, where $S(f)$ and $S(0)$ are the spectral densities at frequency f and 0 Hz respectively; hence $\tau_{noise} = 1/2\pi f_c$ (ms). The single channel conductance, γ , was obtained from both the variance of the current noise and the zero frequency asymptote of the spectra ($\gamma = \sigma^2/\mu_1(V_m - V_{eq})$) and $\gamma = S(0)/2\mu_1\tau(V_m - V_{eq})$, where μ_1 = mean membrane current, V_m = membrane potential, V_{eq} = equilibrium potential for ACh, σ^2 = variance of membrane current fluctuations). In estimating the conductance of the single channel, the equilibrium potential for the endplate current at the human endplate was taken as zero mV (based on unpublished observations).

The mean open time of the ACh-induced channel, obtained by analysis of current noise spectra for normal endplates was, τ_{noise} (normal) = 1.54 ± 0.05 ms (mean $\pm$ s.e.). The comparable value for myasthenic endplates was, τ_{noise} (myasthenic) = 1.63 ± 0.1 ms, at $V_m = -80$ mV and $T = 23^\circ\text{C}$. In addition, the single channel conductance, γ , for normal and myasthenic endplates fell within

the range 20–25 pS. The similarity between the channel characteristics at normal and myasthenic endplates indicates that the properties of ACh-induced channels in myasthenia gravis are virtually unchanged.

As would be expected from earlier work on other mammals, the mean number of transmitter quanta, m , released per nerve impulse from normal and myasthenic nerve-terminals, depends on the extracellular Ca concentration. Figure 2 shows the mean quantal content of e.p.ps, obtained at typical normal and myasthenic endplates, over a range of Ca concentrations. In both cases, m increased as a function of Ca with a slope of approximately 3 on logarithmic co-ordinates. This compares with the value of approximately 3 for rat^{13,14} and for the mouse nerve-muscle junction¹⁵. Although the relation between transmitter release and external Ca concentration had the same slope value for normal and myasthenic nerve terminals, the relationship at myasthenic endplates was shifted to the left when compared with normal. Thus at any given Ca concentration in the range studied (0.3–1.0 mM) more quanta per impulse were released from myasthenic than from normal terminals. For the results from the two endplates illustrated in Fig. 1, the myasthenic terminals released approximately 8 times more quanta per impulse than the normal terminals. The size of this ratio varied from endplate to endplate in different muscles, the mean ratio being around 5:1 (myasthenic: normal) at 26 myasthenic and 12 normal endplates

Fig. 2 Relationship between extracellular calcium concentration and the mean quantal content, m , for a myasthenic endplate (●) and a control endplate (○). Each point represents the average response to 30–180 impulses. Lines with slopes of 3 have been fitted by eye through the points. At any given concentration the myasthenic nerve terminals release approximately eight times more transmitter packets. For the myasthenic curve, m is estimated indirectly from mean e.p.p.²/e.p.p. variance; for the normal curve, m is estimated from mean e.p.p./mean m.e.p.p. at each Ca concentration; additional values at 0.6 and 1.0 mM Ca were estimated from mean e.p.p.²/e.p.p. variance or from $\ln$ (no. of impulses/no. of failures).



investigated over the range of Ca concentrations 0.25–0.7 mM (2 mM Mg). The size of the ratio became less at Ca levels greater than 0.7 mM. In the presence of 2 mM Ca (and 2 mM Mg), myasthenic terminals seem to release about twice as many quanta per nerve impulse as normal terminals. A number of possible factors can be envisaged to account for the higher levels of transmitter release from myasthenic nerve terminals. Briefly, these include: an increase in the synaptic area of myasthenic terminals (ref. 16, but see ref. 17); an increase in the level of intracellular free Ca caused either by a greater than normal influx of Ca during the nerve terminal action potential or by a diminished removal of Ca. Further studies should allow us to distinguish between these and other possibilities.

The release of a larger number of transmitter quanta per nerve impulse, from myasthenic terminals when compared with normal, may be related to biochemical studies which have shown that myasthenic muscle contains about twice as much ACh as normal¹⁸. It may be that transmitter release increases, following the impairment of transmission, as a means of 'compensating' for the reduced postsynaptic sensitivity.

We conclude that the reduction in α -bungarotoxin binding sites and in postsynaptic sensitivity at myasthenic endplates results from a loss of functional receptor-channel complexes rather than a modification of individual ionic channels. In addition, the slope value of about 3 for the Ca-dependence of transmitter release at both myasthenic and normal endplates suggests that the mechanisms underlying Ca-dependence may be unchanged. At a given Ca concentration, however, the number of transmitter packets released per nerve impulse is greater for the myasthenic than for the normal nerve terminals.

We thank Mr J. R. Belcher, F.R.C.S., Dr J. Newsom-Davis, Mr M. F. Sturridge, F.R.C.S. and Dr A. Vincent for their co-operation, and the MRC for support. We are grateful to the patients for their collaboration. S.G.C.C. is a Beit Memorial Research Fellow.

S. G. CULL-CANDY
R. MILEDI
A. TRAUTMANN*

Department of Biophysics,
University College London,
Gower Street, London WC1, UK

Received 18 August; accepted 2 November 1977.

*Present address: Laboratoire de Neurobiologie, Ecole Normale Supérieure, 46 Rue d'Ulm, 75005 Paris, France.

- Grob, D. (ed.) *Myasthenia Gravis* Ann. N.Y. Acad. Sci. **274**, (1976).
- Elmqvist, D., Hofmann, W. W., Kugleberg, J. & Quastel, D. M. J. *J. Physiol., Lond.* **174**, 417–434 (1964).
- Fambrough, D. M., Drachman, D. B. & Satyamurti, S. *Science* **182**, 293–295 (1973).
- Green, D. P. L., Miledi, R., Perez de la Mora, M. & Vincent, A. *Phil. Trans. R. Soc. Lond. B* **270**, 551–552 (1975).
- Albuquerque, E. X., Rash, J. E., Mayer, R. F. & Satterfield, J. R. *Exp. Neurol.* **51**, 536–563 (1976).
- Katz, B. & Miledi, R. *J. Physiol., Lond.* **224**, 665–700 (1972).
- Green, D. P. L., Miledi, R. & Vincent, A. *Proc. R. Soc. Lond. B* **189**, 57–68 (1975).
- Magleby, K. L. & Stevens, C. F. *J. Physiol., Lond.* **223**, 151–171 (1972).
- Gage, P. W. & McBurney, R. N. *J. Physiol., Lond.* **226**, 70–94 (1972).
- Kordas, M. J. *J. Physiol., Lond.* **224**, 317–332 (1972).
- Katz, B. & Miledi, R. *J. Physiol., Lond.* **231**, 549–574 (1973).
- Anderson, C. R. & Stevens, C. F. *J. Physiol., Lond.* **235**, 655–691 (1973).
- Hubbard, J. I., Jones, S. F. & Lundau, E. M. *J. Physiol., Lond.* **196**, 75–86 (1968).
- Cull-Candy, S. G., Lundh, H. & Thesleff, S. J. *J. Physiol., Lond.* **260**, 177–203 (1976).
- Cooke, J. D., Okamoto, K. & Quastel, D. M. J. *J. Physiol., Lond.* **228**, 459–497 (1973).
- Coërs, C. & Telerman-Toppet, N. *Ann. N.Y. Acad. Sci.* **274**, 6–19 (1976).
- Engel, A. & Santa, T. *Ann. N.Y. Acad. Sci.* **183**, 46–63 (1971).
- Ito, Y. *et al. Proc. R. Soc. Lond. B* **192**, 475–480 (1976).

Neural induction of the 16S acetylcholinesterase in muscle cell cultures

SEDIMENTATION analysis of rat muscle acetylcholinesterase (AChE) has revealed the existence of three molecular forms, 4S, 10S and 16S (refs 1, 2). The 16S form was exclusively found in skeletal muscle segments containing neuromuscular junctions; it disappeared after denervation^{1,2} and reappeared following reinnervation either at the original endplate, or at ectopic sites. The 16S AChE may therefore be referred to as an 'endplate specific' form, but it clearly does not represent all activity at the endplate. It took several days for the disappearance of the 16S form after denervation, suggesting a postsynaptic localisation². In posterior leg muscles from rat embryos, the 16S form was absent before the

14th day of gestation; it appeared at day 15 and was relatively more abundant on day 16 (25% of the total AChE activity) than in the adult (5%). (The total AChE activity increased by a factor of 10^3 from day 16 of gestation to the adult muscle².) We present here some data concerning the synthesis of the 16S form by embryonic rat cells in culture, which demonstrate that this form is produced by muscle cell cultures on induction by neuronal elements.

Myoblasts were put in culture after mechanical dissociation of posterior leg muscles, from either 13–14- or from 18-d-old rat embryos. At the earlier age, no 16S AChE could be detected, whereas it was present in the older embryos. Spinal cord cells from

14-d-old embryos were dissociated by gentle trypsinisation or by mechanical techniques, and were plated alone, or added to myoblast cultures. The culture medium was renewed every 3 d for up to 4 weeks, when the cells peeled away from the dishes. The morphology of the cells was examined, and AChE was detected by the histochemical method of Koëlle and Friedenwald³, as modified by Couteaux and Taxi⁴, using acetylthiocholine as substrate. The molecular forms of AChE were analysed, after solubilisation in a buffer containing detergent, by sucrose gradient sedimentation². Their biosynthesis was studied after complete and irreversible inhibition (DFP), by analysing activity during recovery.

Myoblasts from 13–14-d-old embryos possessed only the 4S and 10S forms of AChE, and even after 15 d in culture, it was impossible to detect any 16S form. Similar results were obtained with cultures of spinal cord neurones. When dissociated neurones were added to a myoblast culture (2–3 h after plating, however, we observed the appearance of 16S AChE after 5 d in the mixed cultures (Fig. 1). At this stage, spots of AChE activity were detectable histochemically on the membrane of the myotubes which were coupled with the nerve cells (Fig. 2). Conversely, AChE spots were never detected on myotubes cultivated without nerve cells. In cultures coupled for 10 d, the spots gave place to typical end-plates (J. K., unpublished results). The evolution of the membranous localisation of the AChE activity suggests that the AChE spots observed at 5 d correspond to the foremost neuromuscular contacts. The negative histochemical results of Fischbach *et al.*³ may be due to the fact that the AChE activity of chicken coupled cultures is too weak to be detected by the method they use.

In contrast with 14-d-old myoblasts, muscle cells from 18-d-old embryos contained the 16S form of AChE in addition to the 4S and 10S forms, at the time of plating. The specific activity and the relative proportions of all forms remained stable in the cultures until the cells were peeling from the dishes at 4 weeks. To determine whether this was simply due to the maintenance of AChE molecules, or whether the myotubes were able to synthesise this molecular form, we inhibited all esterase activity in 8-d cultures by treatment with DFP (10^{-5} M for 1 h). Within 48 h, the AChE activity and its distribution in 4S, 10S and 16S forms had recovered. The forms appeared in succession, suggesting that the lighter molecules are precursors of the heavier ones (Fig. 3).

The results presented here clearly demonstrate that the 16S form of AChE may be synthesised by muscle cells, since the contaminating cells of myoblast cultures, fibroblasts and Schwann cells do not contain the 16S form. We never detected this form in cultures of a fibroblast cell line nor in spinal cord cultures which also contain Schwann cells. As mentioned earlier, a muscular

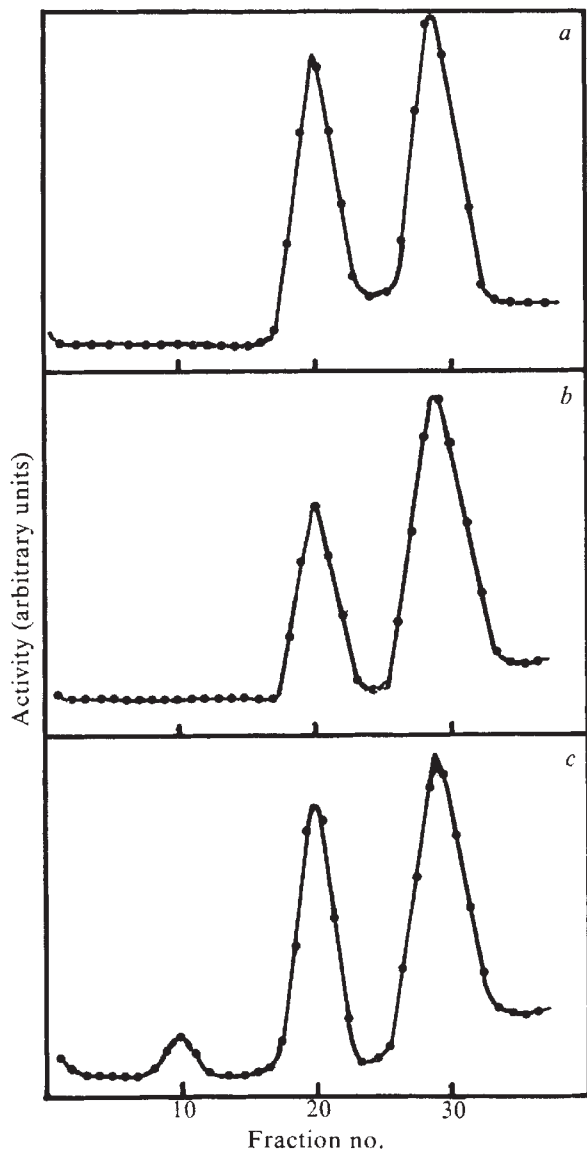
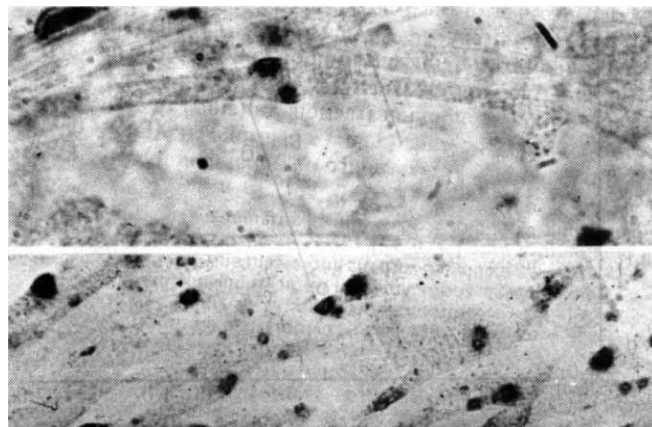


Fig. 1 The molecular forms of acetylcholinesterase in cell cultures of *a*, spinal cord neurones; *b*, myotubes and *c* mixed cultures. 5×10^6 myoblasts or cells from one spinal cord were plated in 50-mm dishes. In mixed cultures, one spinal cord was added to a myoblasts culture 2–3 h later. Cells were grown in 80% MEM, 10% medium 199, 10% horse serum and 0.5% chick embryo extract medium. After 5 d cells were scraped from the dishes, gathered and homogenised using a Potter homogeniser in a solution containing 1 M NaCl, 5×10^{-2} M $MgCl_2$, 10^{-2} M Tris-HCl pH 7 and 1% Triton X-100. The homogenate was centrifuged at 20,000g for 30 min at 4 °C and the supernatant assayed immediately for AChE activity and analysed on a sucrose gradient. AChE activity was determined at 20 °C by the method of Ellman *et al.*⁵. Sedimentation ultracentrifugation on linear sucrose gradients (5–20%) in presence of the homogenising buffer was performed on a Spinco LII 65 ultracentrifuge, using a SW60 rotor at 125,000g for 15 h at 2 °C. Fractions were collected and assayed immediately for enzymatic activities. B-Galactosidase (16S) and horse liver alcohol dehydrogenase (4.8S) were used as enzyme markers. The recovery of AChE activity after sedimentation procedures was higher than 80%.

Fig. 2 Areas of high activity of AChE visualised as 'spots' on the membranes of young myotubes by the histochemical method of Koëlle and Friedenwald⁶ as modified by Couteaux and Taxi⁷. Fixation was in neutral formaldehyde 5% for 5 min, followed by incubation in Koëlle's medium using AChE as substrate at pH 5 for 12 h. Cultures of myoblasts obtained from 13–14-d-old embryos coupled for 5 d with cord cells of 14-d-old embryos. The cell culture conditions were as in Fig. 1 $\times 800$.



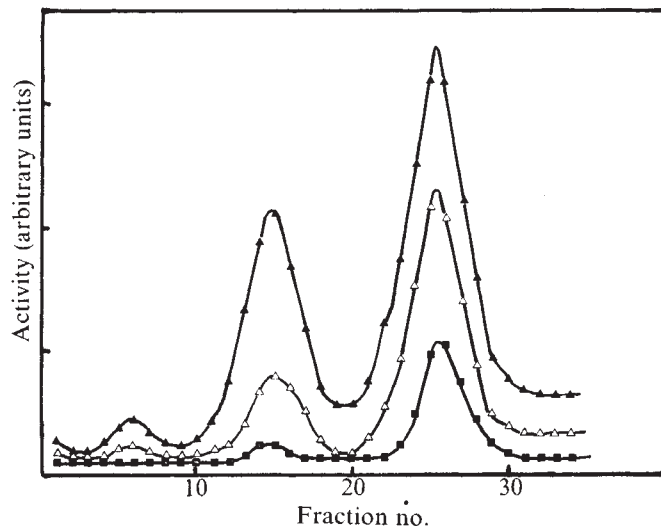


Fig. 3 Restoration of AChE activity after DFP inhibition in cell culture of myotubes from 18-d-old embryos. Here 10^6 myoblasts were plated in 60-mm dishes. Cells were incubated for 1 h in presence of 10^{-5} M DFP in the culture medium to inhibit AChE activity. Such treated cells were washed in order to eliminate excess free DFP and incubated in the culture medium for periods indicated. Other experimental conditions were as in Fig. 1. ■, Cells 7 h after DFP incubation; Δ, cells 14 h after DFP incubation; ▲, cells not incubated with DFP or cells 48 h after DFP incubation.

localisation was implied by the slowness of its disappearance in denervation experiments, but this was not conclusive for its biosynthesis.

From studies of neuroblastoma cell cultures⁴ it has been concluded that the 4S form is a precursor of the 10S form. The present observations also suggest that these forms might in turn be precursors of the 16S molecule.

The most interesting feature of the 16S form of AChE is that its synthesis by muscle cells requires an interaction with neuronal elements, since it only appears in mixed cultures, or in cultures from 18-d-old embryo cells, which have made contact with neurones *in vivo* before being plated. In this second case, the continued synthesis of 16S AChE by these cells, in the absence of neurones, seems paradoxical in view of the disappearance of this form from denervated adult muscles. The experimental situations are, however, so different that it is difficult to interpret this phenomenon and it might well reflect the plasticity of embryonic nerve-muscle interactions.

It would be extremely useful to understand how the 16S molecule is constructed and if non-catalytic elements are included in this complex structure, in addition to knowing how their synthesis and assembly are regulated. The observations reported here not only demonstrate the muscular origin of the 16S acetylcholinesterase and its induction by neuronal elements, but the two types of cultures described provide excellent models for future investigation of the nature of the interaction involved.

We thank Professor G. Buttin and Dr M. Lambiotte for culture facilities and Ms P. Cloup, M. Orouet and D. Duflou for technical assistance. We are grateful to Dr J. Massoulié for helpful discussions and critical comments. This work was supported by funds from the Centre National de la Recherche Scientifique and from the Institut National de la Santé et de la Recherche Médicale.

J. KOENIG

Université P. et M. Curie,
Laboratoire de Cytologie,
7, quai St Bernard, 75005 Paris and
E.R. 10007, CNRS,
Hôpital de la Salpêtrière,
75013 Paris, France

M. VIGNY

Laboratoire de Neurobiologie,
Ecole Normale Supérieure,
46, rue d'Ulm, 75005 Paris, France

- Hall, Z. J. *Neurobiol.* **4**, 343-361 (1973).
- Vigny, M., Koenig, J. & Rieger, F. J. *Neurochem.* **27**, 1347-1353 (1976).
- Fischbach, G. D., Berg, D. K., Cohen, S. A. & Frank, E. *Cold Spring Harbor Symp. quant. Biol.* **40**, 347-357 (1976).
- Rieger, F., Bauman, A. F., Benda, P. & Vigny, M. J. *Neurochem.* **27**, 1059-1063 (1976).
- Ellman, G. L., Courtney, K. D., Andres, V. & Featherstone, R. M. *Biochem. Pharmacol.* **7**, 89-95 (1961).
- Koelle, G. B. & Friedenwald, J. S. *Proc. Soc. exp. Biol. Med.* **70**, 617-622 (1949).
- Couteaux, R. & Taxi, J. *Archs Anat. microsc. Morph. exp.* **41**, 452-492 (1952).

Effects of $[Ca^{2+}]$ and $[Mg^{2+}]$ on the decay of miniature endplate currents

THE endplate current at the neuromuscular junction is initiated by the combination of acetylcholine (ACh) with its receptor. The current rises rapidly to a maximum, and then decays exponentially¹. The rate of decay of the endplate current (e.p.c.) depends on membrane potential; hyperpolarisation lengthens the decay phase^{2,3}. A plausible hypothesis for this is that the rate of disassociation of the ACh from the receptor depends on the voltage gradient through the membrane⁴. The membrane potential consists of three components: the voltage gradient through the membrane itself and the surface potentials at the two membrane-solution interfaces. The surface potentials are thought to be negative, owing to an excess of negatively-charged groups on the membrane phospholipids. An elevation in the divalent ion concentration in the extracellular solutions reduces the negativity of the external surface potential, either by binding to the phospholipids or by screening the charges. A lowering of the negativity of the external surface potential leads to a corresponding increase in the voltage gradient within the membrane (over shorter periods than for phospholipid flip-flop to occur). This leads to the prediction that alterations in the extracellular $[Ca^{2+}]$ or $[Mg^{2+}]$ are likely to change the time course of the e.p.c. Our experiments were designed to test this hypothesis.

We studied spontaneously occurring miniature endplate currents (m.e.p.cs) to avoid problems caused by raised divalent ion concentrations, such as, changing quantal outputs from stimulated nerves or alterations in the shape of the presynaptic action potential. The two-microelectrode voltage-clamp technique was used at the neuromuscular junction of the sartorius muscle of the frog, *Rana pipiens*¹. The microelectrodes were inserted close to the centre of the junction, as judged by the rise times of the miniature endplate potentials (m.e.p.ps). One electrode monitored the membrane potential, the other passed the clamping current. Both were filled with 4 M potassium acetate. Experiments were carried out both in the presence and absence of 10^{-6} g ml⁻¹ neostigmine methylsulphate, which has been shown to slow the decay phase of the m.e.p.cs without altering their voltage dependence⁴.

The results of one experiment in the presence of neostigmine are shown in Fig. 1. The $[Ca^{2+}]$ concentration was varied tenfold and m.e.p.cs were recorded between -83 mV and -150 mV. From the recordings, decay half-times ($\tau_{1/2}$) were measured. The tenfold increase in $[Ca^{2+}]$ caused roughly the same slowing of the decay as a hyperpolarisation of 20 mV. In six additional experiments, four in the presence of neostigmine and two in its absence, there was a consistent slowing of the decay time course of the m.e.p.cs in elevated $[Ca^{2+}]$. (The magnitude of this effect seemed somewhat smaller in the neostigmine-free preparations.)

To test whether this effect was specific to Ca^{2+} or a more general characteristic of divalent cations, the experiments were repeated varying the Mg^{2+} concentration. One example in the absence of neostigmine is shown in Fig. 2. In this experiment there was a slowing of the decay of the m.e.p.cs equivalent to a 12-mV hyperpolarisation when the $[Mg^{2+}]$ was altered from 1 to 10 mM. Similar effects were seen in six additional experiments (two more in the absence of neostigmine and four in its presence).

Since Mg^{2+} and Ca^{2+} produce similar effects in lengthening the decay phase of the m.e.p.cs it seemed possible that the effects might be produced exclusively by the screening of the surface charges by the divalent ions in the solution. The relation between membrane surface charge density, σ (charges cm⁻²), and membrane surface potential, ψ (V), in solutions with different con-

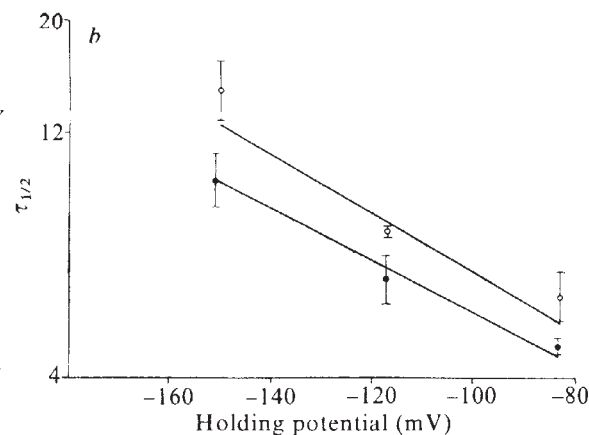
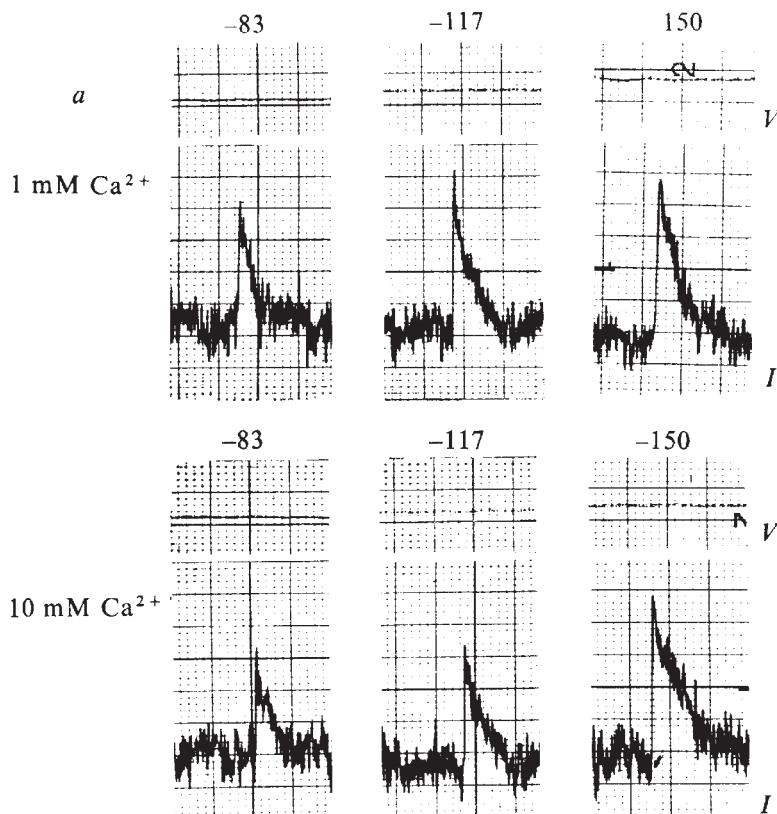


Fig. 1 *a*, Examples of the effects of changing membrane potential and $[Ca^{2+}]$ on m.e.p.s (I). The upper line in each frame (V) shows the membrane potential, illustrating the effectiveness of the clamp. Note that a hyperpolarisation slows the rate of decay of the m.e.p.s. Calibrations: one large vertical box is 0.1 mV, in the upper line, and 1 nA in the lower line of each frame; one large horizontal box is 20 ms. *b*, The half-time for the decay of the m.e.p.s as a function of the membrane holding potential in 1 mM (●) and in 10 mM (○) Ca^{2+} . Bars represent $\pm$ s.e.m. The Ringer contained (in mM) 115 NaCl, 2 KCl, 4 N-Tris-(hydroxymethyl) methyl-z-aminothanesulphonic acid.

concentrations of univalent, i^+ , and of divalent, i^{2+} , ions (mol cm^{-3}) is

$$\sigma^2 = \frac{\epsilon k T}{2\pi} \left[[i^+] [\exp(Q\psi/kT) + \exp(-Q\psi/kT) - 2] + [i^{2+}] [2\exp(Q\psi/kT) + \exp(-2Q\psi/kT) - 3] \right] \quad (1)$$

where k is the Boltzmann constant, T is temperature in degrees Kelvin, Q is the charge on the electron, and ϵ is the permittivity of water⁵.

The data in Fig. 1 show that a tenfold increase in $[Ca^{2+}]$ produces an effect that could be attributed to a 20-mV decrease in ψ . By setting up equation (1) for the two divalent ion concentrations, it is possible to solve the pair of simultaneous equations for ψ . In 1 mM $[Ca^{2+}]$, ψ is about -100 mV, which would be produced by a charge density, σ , of about 1 charge per $93 \text{ \AA}^2$. Such a high surface potential seems unlikely and in other experiments the change in ψ produced by a tenfold rise in extracellular divalent ion concentration seems to be even greater. Consequently it is unlikely that the changes in ψ could be produced by screening alone, there is probably also some binding of divalent ions to the membrane. If we assume that ψ in 1 mM Ca^{2+} is -75 mV ($\sigma = 1$ charge per $181 \text{ \AA}^2$), a 20-mV fall in ψ would follow a 20% decrease in the number of excess negative charges on the membrane surface (to $\sigma = 1$ charge per $302 \text{ \AA}^2$).

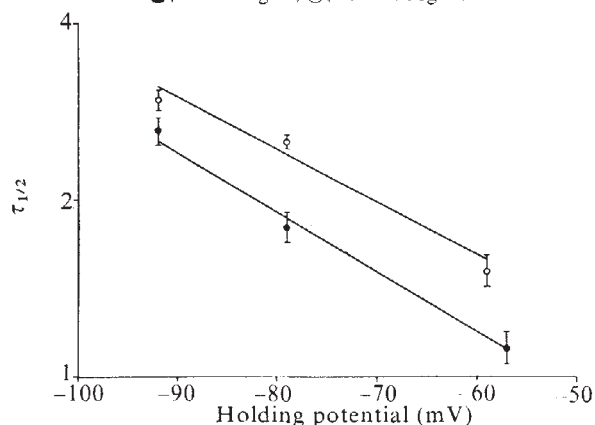
Any alteration of the surface potential will change the concentration profiles of ions between the bulk solutions on the two sides of the membrane, and thus alter the single channel conductance. In all experiments in 10 mM Mg^{2+} and Ca^{2+} we noticed a decrease in the average peak height of m.e.p.s. This result agrees with experiments showing, by noise analysis, a decrease in the single channel conductance at increased $[Ca^{2+}]$. (C. Lewis and C. F. Stevens, personal communication.)

Inspection of Figs 1 and 2, together with data not reproduced, suggests that the lines connecting the data points are not parallel, they often seem to converge or diverge at the most negative holding potentials. This could be an artefact caused by the distribution of the endplate, which extends for an appreciable fraction of a length constant along the fibre. Therefore, when the

fibre is clamped away from the resting potential, there is a voltage gradient between the current-passing electrodes and the margins of the endplate. If the resting potential falls, as is almost inevitable following repeated periods of current injection, then the voltage gradient increases, and this may lead to errors in relating the observed decay half-times to the potential set by the clamp near the centre of the endplate.

In summary, the results are consistent with the idea that the membrane at the muscle endplate has a net negative external surface charge. The lowering of the surface potential by screening and/or binding of divalent ions increases the voltage gradient within the membrane. The increased voltage gradient within the membrane slows the time course of the decline of the ACh-induced conductance change. These results have implications for other work. For example, experiments in which the Ca^{2+} and Mg^{2+} levels are manipulated to change the quantal outputs from the stimulated nerve terminals will also be changing the time course of the post-junctional current flows, unless the total concentrations of divalent cations are maintained at a constant

Fig. 2 Effects of changing membrane potential and $[Mg^{2+}]$ on the $\tau_{1/2}$ of the m.e.p.s. Solutions are as in Fig. 1b with Mg^{2+} substituted for Ca^{2+} . Bars represent $\pm$ s.e.m. ●, 1 mM Mg^{2+} ; ○, 10 mM Mg^{2+} .



level. The post-junctional effects must be considered carefully when evaluating the results. The effects of negative surface charges near the ACh receptor will change the local concentrations of all charged substances near the surface of the membrane according to the Boltzmann distribution:

$$[i^N]_0 = [i^N] \exp[-NQ\psi/kT] \quad (2)$$

where $[i^N]$ is the concentration of I in the bulk solution $[i^N]_0$, the concentration at the membrane surface, and N is the valance. For example, if the ψ at the outer membrane surface is -70 mV, a positive univalent ion, like ACh, will be 15.4 times more concentrated at the surface. A positive divalent ion, like tubocurarine, will be concentrated more than 200-fold.

This study was supported by a grant from the NINCDS, and by an Institutional Research Support Grant from the US Public Health Service. The authors thank Mr David Attwell for discussion.

IRA COHEN
WILLIAM VAN DER KLOOT

Department of Physiology and Biophysics,
Health Sciences Center,
State University of New York at Stony Brook,
Stony Brook, New York 11794

Received 26 September; accepted 25 October 1977.

1. Takeuchi, A. & Takeuchi, N. *J. Neurophysiol.* **22**, 395-411 (1959).
2. Gage, P. W. & Armstrong, P. W. *Nature* **218**, 363-365 (1968).
3. Kordas, M. J. *Physiol. Lond.* **204**, 493-502 (1969).
4. Magleby, K. L. & Stevens, C. F. *J. Physiol. Lond.* **223**, 151-171 (1972).
5. Grünhe, D. S. *Chen. Rev.* **41**, 441-501 (1947).

Phospholipase A and the mechanism of action of aldosterone

ALDOSTERONE stimulates Na transport across the renal tubule and this effect is also seen in 'model' epithelia such as amphibian urinary bladder, skin and colon¹⁻³. This hormone is thought to act by inducing the formation of proteins through nuclear transcription⁴. Their nature and actions are unknown but it has been suggested that one such protein may act like a 'permease' and increase the permeability of the apical side of the epithelial cells or they could stimulate the activity of the Na transport pump. Goodman *et al.*⁵, have suggested that "some key enzyme" which increases the turnover of fatty acids may be involved and this could be phospholipase. We report here the identification of phospholipase A in toad bladder and frog skin epithelial cells and in their incubation media. Mepacrine, an inhibitor of phospholipase A, blocked the response to aldosterone on Na transport.

Toad urinary bladder and frog skin were used. Phospholipase activity was determined in tissue extracts and fluids using phosphatidyl[N-methyl¹⁴C]choline 60mCi mmol⁻¹ (Amersham-Searle) as a substrate (Table 1). Phospholipase activity was determined in tissue or cell homogenates by incubation, in 2-4 ml of Ringer solution containing 2-3 nmol of ¹⁴C-phosphatidylcholine for 2 h at 22 °C. Enzyme activity in the fluids bathing the frog skin and toad bladder preparations was determined by incubating the tissues in Ringer solution and after a specified time removing 4 ml of fluid from both sides, and incubating this with the labelled substrate. The enzyme reaction was terminated by addition of a 2:1 chloroform-methanol mixture and 1 ml of the aqueous phase was counted in a liquid scintillation counter.

Homogenates of the toad bladder and frog skin had phospholipase activity and this was localised in the separated epithelial cells (Table 1). No significant activity

was found in the liver or skeletal muscle of the toads. Gentle homogenisation of the isolated epithelial cells resulted in partial release of the enzyme in the first supernatant fraction (Table 1). The remaining activity appeared in the aqueous following sonication of the pellet. If, however, the tissues were initially sonicated, nearly all the activity could be identified in the supernatant separated by high speed centrifugation. These results suggest that while the activity may be sequestered, it is readily released, and is soluble.

We incubated bladder and skin preparations in Ringer and assayed the fluid that bathed each surface (Table 2). These solutions could also hydrolyse ¹⁴C-lecithin and the total amount of activity that was released in 3 h far exceeded that stored in the tissues. However, it is possible that endogenous unlabelled lecithin could be interfering with the tissue assays, resulting in an underestimate, though this seems unlikely in the supernatants separated by high speed ultracentrifugation. We also identified the enzyme material in the mucosal fluid bathing the toad's colon (which transports Na *in vitro*³) (Table 2), and in its urine (1 ml (14 samples) hydrolysed 90 ± 14 p mol ¹⁴C-lecithin⁻¹). Some bacteria form phospholipases so that we also included either gentamicin (Schering; 100 µg ml⁻¹) or penicillin (Pfizer, 100 U ml⁻¹) plus streptomycin (Pfizer, 100 µg ml⁻¹) in the bathing and incubation media, but they did not decrease the accumulated enzyme activity. In the bladder the rate of secretion was asymmetrical, more appearing in the mucosal than serosal fluid. In the skin a consistent pattern was not seen.

The products of the reaction of the enzyme secreted by the toad bladder epithelial cells were identified using thin

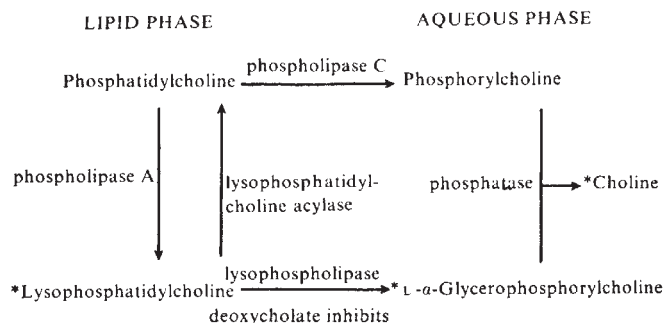


Fig. 1 The possible pathways for phosphatidylcholine metabolism and the products (*) which were identified in response to the action of the phospholipase secreted by amphibian epithelia. To identify the products of phospholipase action the lipid and aqueous phases were each subjected to thin layer chromatography. The two phases were separated and taken to dryness under reduced pressure. The remaining material was dissolved in 200 µl of methanol and 20 µl of this solution was applied to thin layer plates. Chromatographic separation of the phospholipids (lipid phase) was carried out in a solvent mixture consisting of chloroform, methanol, water (65:25:4) on silica gel G plates. The fractions were visualised by exposure to iodine vapour and identified by comparing them to appropriate reference substances. After the disappearance of the iodine staining the areas of the thin layer corresponding to phosphatidylcholine and lysophosphatidylcholine were scraped and eluted with methanol directly into vials and counted in the liquid scintillation spectrometer. The quantity of silica gel of each scraped area was constant. The relative mobilities (R_f) of these phospholipids in the solvent system used were: phosphatidylcholine, 0.34; lysophosphatidylcholine, 0.15. The labelled substrate, ¹⁴C-phosphatidylcholine, when chromatographed on thin layer plates alone was found to be >99% pure. The water soluble products were chromatographed on cellulose sheets (Kodak) using a solvent system of chloroform, methanol, ammonia (65:45:8) and visualised by spraying with a modified molybdate solution and exposing the plates to ultraviolet light for 30 min. The areas corresponding to phosphorylcholine, L-α-glycerophosphorylcholine and choline were scraped off and eluted in counting vials as described above. The R_f values were: phosphorylcholine, 0.02; L-α-glycerophosphorylcholine, 0.13 and choline, 0.40.

layer chromatography (Fig. 1). Small amounts of choline were present but the principal component in the aqueous fraction behaved chromatographically like 1- α -glycerophosphorylcholine. The formation of this lipid could be prevented by the presence of 5 mM deoxycholate, which inhibits lysolecithinase⁹. Lysophosphatidylcholine (lysolecithin) accumulated in the lipid phase in the presence of deoxycholate but could not be detected in its absence. Thus the primary product of the enzymatic reaction is apparently lysophosphatidylcholine, which is formed as a result of the action of phospholipase A on phosphatidylcholine, but this is apparently normally converted to water soluble 1- α -glycerophosphorylcholine. The 'lecithin-lysolecithin cycle' has previously been identified in homogenates of toad bladder⁷.

It has been suggested that a phospholipase may be involved in the action of aldosterone on epithelial membranes³. This could involve an increased rate of turnover of the enzyme or an increase in its activity. The demonstration of such changes may be expected to be a problem as the enzyme is readily inactivated. Thus in the bladder the secreted activity less than doubles between 3 and 18 h incubation (Table 2). Indeed, we were unable to demonstrate a change in secretion of the enzyme in response to aldosterone in the bladder. However, in another target tissue, the frog skin, accumulation increased fourfold. Frog skin moults in the presence of aldosterone, though both responses are superimposed and involve the same types of cell⁸. The morphological change may favour the release or stability of the enzyme.

Phospholipase A is inhibited by mepacrine^{9,10}; at 10⁻⁶ M mepacrine inhibited the *in vitro* activity of the enzyme secreted by the toad bladder by 35%. Mepacrine also completely inhibited the effects of aldosterone on Na

Table 1 Phospholipase activity in isolated epithelial cells from frog (*Rana pipiens*) skin and toad (*Bufo marinus*) urinary bladder

	¹⁴ C-phosphatidylcholine hydrolysed (pmol h ⁻¹)
Epithelial cells, toad bladder	
a, Whole homogenate, sonicated (8)	64 ± 8
Boiled (6)	7 ± 0.7
b, 'Mild' homogenisation and extraction (6)	
Supernatant (150,000g for 1 h)	71 ± 8
Pellet sonicated	
Supernatant (150,000g for 1 h)	67 ± 11
Pellet suspension	32 ± 9
c, Cells immediately homogenised and sonicated (6)	
Supernatant (150,000g for 1 h)	138 ± 21
Pellet resuspended	
Supernatant (150,000g for 1 h)	14 ± 3
Pellet suspension	17 ± 5
Isolated epithelium, frog skin	
Whole homogenate, sonicated (6)	60 ± 9
Boiled (5)	11 ± 1.6

The toads weighed about 200 g. The bladders (two hemi-bladders, about 200 mg whole tissue) were perfused with Ringer to remove the blood. The bladder epithelial cells were separated by incubating each hemi-bladder sac preparation in Ringer (in mM, NaCl, 111; KCl, 3.35; CaCl₂, 2.7; NaHCO₃, 4.0 and glucose 5.0) in the absence of added Ca²⁺ and containing 2 mM EGTA, for 75 min. The cells were separated by gently massaging the sac, pooled and centrifuged for 15 min at 2,500g. Pieces of frog skin epithelium were separated by exposing the inner surface to collagenase (Sigma, 250 µg ml⁻¹) under about 50 cm hydrostatic pressure for about 2 h. The isolated bladder epithelial cells were resuspended in 2 ml of a solution containing 4 mM NaHCO₃ and 5 mM CaCl₂ (pH about 8.0), and incubated with 2-3 nmol of phosphatidyl[N³-methyl-¹⁴C]choline for 2 h at 22 °C. The solution was then extracted with 19 ml of 2:1 chloroform-methanol and 1 ml of the aqueous was counted in a scintillation counter. Assays were corrected for a reaction 'blank' which was small and amounted to about 0.4% of the total phosphatidylcholine present. Frog skin epithelium (20 cm², 424 mg whole skin) was assayed similarly. Results are means ± s.e. Numbers of experiments carried out shown in parentheses.

Table 2 Secretion of phospholipase by frog skin and toad bladder

	¹⁴ C-phosphatidylcholine hydrolysed (pmol h ⁻¹)
Toad bladder: Mucosal side 3 h (8)	113 ± 20
Serosal side 3 h (6)	46 ± 11
Mucosal side 3 h, boiled (6)	5 ± 1
Saline treated toads:	
Total (mucosal and serosal) 3 h (8)	58 ± 18
Frog skin: Epidermal side 3 h (10)	64 ± 10
Corium side 3 h (10)	67 ± 17
Toad colon: mucosal side 3 h (5)	39 ± 6
Effect of 10 ⁻⁶ M aldosterone; (total secretion)	
Toad bladder; 18 h incubation*	
Control (6)	260 ± 55
Aldosterone (6)	274 ± 60
Frog skin; 18 h incubation	
Control (10)	318 ± 25
Aldosterone (10)	1,267 ± 198†

Toad hemi-bladders (and colons) were prepared as sacs. The short-circuit current (µA per 100 mg) for the 18 h incubation was 158 ± 20 for the control, and 555 ± 62 for aldosterone. The frog skins were either tied onto lucite cylinders (surface area 10 cm² - 200 mg) or for testing the effects of aldosterone, prepared in Ussing-type chambers (3 cm² = 60 mg). The short-circuit current after 18 h was (µA cm⁻²) 17 ± 2 in the control and 42 ± 8 in the aldosterone treated skin. Values are mean ± s.e., pmol hydrolysed in fluids bathing 100 mg tissue. Number of experiments shown in parentheses.

*We were also unable to see a stimulation of enzyme activity by aldosterone even for the shorter time period of 3 h either initially or following an 18 h preincubation.

†P < 0.01 for difference from control.

transport in the toad bladder (Table 3) though it had no significant effect on the natriiferic response to vasopressin (ADH). Phospholipase A₂ is the rate limiting step in the formation of prostaglandins. This enzyme releases free

Table 3 Effects of inhibitors of phospholipase A and prostaglandin synthetase on the response of the toad bladder (*in vitro*) to aldosterone (10⁻⁷ M) and vasopressin (AVP, 10 mU ml⁻¹)

	Short circuit current (µA per 100 mg wet wt)		
	Initial	Final 5 h	Mean difference
I a, Control (6)	276	251	-25 ± 24
+ 10 ⁻⁶ M Mepacrine	229	219	-10 ± 28
b, Aldosterone (6)	345	762	434 ± 118*
+ 10 ⁻⁶ M Mepacrine	357	452	96 ± 49
c, Aldosterone (6)	318	764	446 ± 127*
+ 5 × 10 ⁻⁶ M Mepacrine	424	334	-90 ± 46
II a, Control (12)	168	198	30 ± 16
+ 10 ⁻⁵ M Indomethacin	207	98	-109 ± 22†
b, Aldosterone (6)	295	775	480 ± 120*
+ 10 ⁻⁵ M Indomethacin	350	439	89 ± 46
III a, Control (12)	234	252	18 ± 30
+ 10 ⁻⁶ M Mefenamic acid	244	159	85 ± 42
b, Aldosterone (6)	231	681	450 ± 111†
+ 10 ⁻⁶ M Mefenamic acid	291	437	146 ± 62
	Initial	30 min	Mean difference
IV a, AVP (6)	251	621	370 ± 34†
+ 10 ⁻⁶ M Mepacrine	219	582	363 ± 18†
b, AVP (12)	198	715	517 ± 106†
+ 10 ⁻⁵ M Indomethacin	98	548	450 ± 74†
c, AVP (12)	252	820	568 ± 119†
+ 10 ⁻⁵ M Mefenamic acid	159	588	429 ± 69†

Toad hemi-bladders were prepared as sacs. Paired bladder lobes were used. The number of paired experiments are in parentheses. Short circuit current reflects Na transport. In IV the tissues were incubated for 5 h in the presence or absence of the drug, before adding the AVP.

*P < 0.02 for differences from initial period.

†P < 0.01.

fatty acids which are substrates for prostaglandin synthesis. It was therefore also interesting to observe that mefenamic acid and indomethacin which inhibit prostaglandin synthetase¹¹ also inhibited the action of aldosterone on Na transport, though the response to vasopressin was unaffected (Table 3).

The effect of aldosterone on transepithelial Na transport thus seems to be related to the actions of two enzymes, phospholipase A and prostaglandin synthetase. Both of these are known to be present in this tissue and are known to be involved in the common chain of events that mediates (this paper and ref. 12) prostaglandin synthesis. The mechanism of aldosterone action is unclear, but it could be inducing the formation, activation, or release of phospholipase A. It has recently been shown that RNA and protein synthesis are necessary for expression of phospholipase action in mouse fibroblasts¹³. We do not know by which process prostaglandins could be influencing the response but some are known to increase active Na transport across these membranes^{14,15}.

This work was supported by the US NSF. Dr L. A. Barker kindly helped and advised us in the identification of phospholipase A.

THOMAS YORIO

P. J. BENTLEY

Departments of Pharmacology and Ophthalmology,
Mount Sinai School of Medicine,
New York, New York 10029

Received 8 July; accepted 11 November 1977.

- Crabbé, J. *Endocrinology* **69**, 673-682 (1961).
- Crabbé, J. & DeWeert, P. *Nature* **202**, 298-299 (1964).
- Cofré, G. & Crabbé, J. *Nature* **207**, 1299-1300 (1965).
- Feldman, D., Funder, J. W. & Edelman, I. S. *Am. J. Med.* **53**, 545-560 (1972).
- Goodman, D. B. P., Wong, M. & Rasmussen, H. *Biochemistry* **14**, 2803-2809 (1975).
- Elsbach, P., Van Den Berg, J. W. O., Van Den Bosh, H. & Van Deenen, L. L. M. *Biochim. biophys. Acta* **106**, 338-347 (1965).
- Rosenbloom, A. A. & Elsbach, P. *Am. J. Physiol.* **214**, 161-165 (1968).
- Voute, C. L., Hanni, S. & Ammann, E. J. *Steroid Biochem.* **3**, 161-165 (1972).
- Flower, R. J. & Blackwell, G. J. *Biochem. Pharmacol.* **25**, 285-291 (1976).
- Vargaftig, B. B. *J. Pharm. Pharmacol.* **29**, 222-228 (1977).
- Flower, R. J. & Vane, J. R. *Biochem. Pharmacol.* **23**, 1439-1450 (1974).
- Wong, P. Y. D., Bedwani, J. R. & Cuthbert, A. W. *Nature* **238**, 25-31 (1972).
- Pong, S.-S., Hong, S. L. & Levine, L. J. *J. Biol. Chem.* **252**, 1408-1413 (1977).
- Lipson, L. C. & Sharp, G. W. G. *Am. J. Physiol.* **220**, 1046-1052 (1971).
- Barry, E. & Hill, W. J. *J. Physiol. Lond.* **200**, 83-84 (1969).

Inactivation of interferon mRNA in the shutoff of human interferon synthesis

INDUCTION of interferon synthesis in cultured cells by $rI_n \cdot rC_n$ has been intensely studied because of its attractiveness as a model for gene expression in mammalian cells and its possible significance to medicine. Exposure of human fibroblast FS-4 cells to $rI_n \cdot rC_n$ results in interferon (IF) synthesis which persists for 6 h and then abruptly shuts off^{1,2}. The rapidity of this shutoff indicates that the cessation of IF mRNA transcription is unlikely to be the primary control. Moreover, studies with metabolic inhibitors³⁻⁷ have suggested that a post-transcriptional event is involved. To evaluate the role of post-transcriptional control in the shutoff of IF production, we measured directly the level of IF mRNA activity in a strain of human fibroblast HF 926 cells during the IF induction process. This was achieved by the development of a highly efficient and reproducible heterologous whole cell translational system. We report here that, using this system, we have found a significant decrease in the IF mRNA activity during the shutoff of IF protein synthesis.

The translational system consists of the direct application of isolated human RNA to Syrian hamster embryo (SHE) cells. Absolute species specificity between the human and hamster interferons allows for the bioassay of human IF

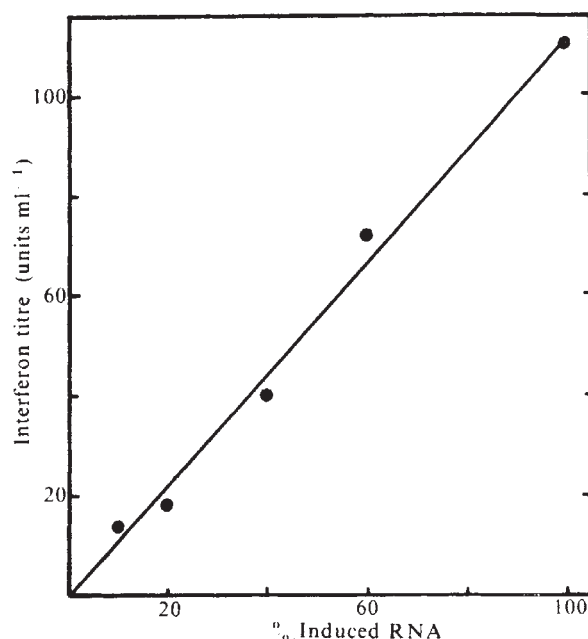
Table 1 Effect of actinomycin-D pre-incubation and RNA facilitators on interferon mRNA Translational efficiency

Treatment	Incubation time (h)	Facilitator	IF titre (units ml ⁻¹)	% Efficiency
None	-	CaCl ₂	13	6.25
Actinomycin D	5	CaCl ₂	22	11.0
Actinomycin D	5	DEAE-Dextran	4	2.0
Actinomycin D	10	CaCl ₂	22	11.0
Actinomycin D	15	CaCl ₂	41	20.0

Confluent monolayers of Syrian hamster embryo cells (SHE) in 75 cm² flasks with $3-4 \times 10^6$ cells per flask were treated with 0.065 μ g ml⁻¹ actinomycin D in maintenance media for the times indicated. After actinomycin treatment, 50 μ g ml⁻¹ cytoplasmic RNA, obtained from $rI_n \cdot rC_n$ -induced HF-926 cells by a modified Penman procedure¹³, were applied to the SHE cells with either 50 μ g ml⁻¹ of DEAE-dextran or 30 mM CaCl₂ in 5 ml of HEPES-phosphate-buffered saline (PBS) buffer, pH 7.0. After incubating at 37 °C for 1 h, the HEPES-PBS buffer was removed from all flasks and replaced with 10 ml of maintenance media. Translation was allowed to proceed for 5 h and the media was collected for assay.

that has been translated in the heterologous SHE cells. Translation is dependent on application of the RNA with a substance which facilitates the RNA uptake. Both DEAE-dextran⁸ and CaCl₂ (refs 9, 10) have been used successfully to enhance the infectivity of viral nucleic acids. We adapted these procedures for mRNA. CaCl₂ was a more effective facilitator than DEAE-dextran in both translational efficiency and reproducibility (Table 1). Pretreatment of the recipient SHE cells with actinomycin-D before addition of human IF mRNA was not essential, but this treatment increased translational efficiency, particularly with long preincubation times. Presumably, this enhancement is the consequence of endogenous mRNA degradation in the SHE cells with the resultant increase in translational capacity. Efficiency of translation can be estimated by comparing the amount of human IF translated by the SHE cells from a given quantity of added human RNA to the amount of IF

Fig. 1 Cytoplasmic RNA extracted by a modified Penman procedure¹³ from HF 926 cells 3 h after induction with $rI_n \cdot rC_n$ was mixed with cytoplasmic RNA from mock induced HF 926 cells at the shown ratios. The total concentration of RNA in each sample was 50 μ g ml⁻¹. All points represent the average of duplicate assays.



produced from an equivalent number of human cells from which the RNA was isolated (Table 1). For instance, in a typical experiment, one 1/2 gallon roller bottle (690 cm², 2–3 × 10⁷ cells) of induced human cells produced approximately 10,000 IF units, whereas approximately 600 IF units (6% efficiency) were produced by the SHE cells on translation of the mRNA isolated from this roller bottle. Depleting the SHE cells of Ca²⁺ before addition of human RNA also increases IF mRNA translation (Table 2). Interferon activity was always associated with the poly A mRNA fraction retained on oligo dT-cellulose and not with the flow-through fraction. No IF production was found when RNA from uninduced human cells was used in translation.

When cytoplasmic RNA extracted from rI_h-rC_h-induced HF 926 cells was applied to Ca²⁺-depleted SHE cells using the CaCl₂ precipitation technique, the IF translation was found to be independent of the absolute RNA concentration. In the concentration range of 10 µg ml⁻¹ to 100 µg ml⁻¹, the IF yield remained constant. But, when RNA from induced cells was diluted with varying amounts of carrier RNA while keeping the total RNA concentration constant at 50 µg ml⁻¹, IF production varied linearly with the proportion of IF mRNA in the applied RNA sample (Fig. 1). These observations can be explained on the basis of a limitation in the RNA uptake of SHE cells. If the uptake capacity is limited, then the absolute amount of mRNA translated by the SHE cells is fixed when the amount of applied RNA exceeds this capacity. Therefore, in saturating conditions (over 10 µg RNA per ml) the SHE cell translational system responds proportionately not to absolute dosage of IF mRNA but rather to its relative concentration in the preparation (equivalent to 'specific biological activity' of IF mRNA).

Kinetics of the IF protein synthesis measured in HF 926 cells after 1 h induction with 0.4 mM (130 µg ml⁻¹) rI_h-rC_h in physiological salt buffer (0.15 M NaCl, 0.01 M PO₄, 0.001 M MgCl₂ pH 7.2), is shown in Fig. 2a. The rate of IF synthesis reached a maximum at about 3–4 h after induction and was largely shut off by 8 h. In many of these experiments, especially when a lower concentration of rI_h-rC_h was used for induction, there was no detectable synthesis at 8 h. These findings corroborate those of Vilček and his colleagues¹¹ for FS-4 cells.

Total cytoplasmic RNA was extracted according to a modified Penman procedure^{12,13} from the same cells and at the same time points as in Fig. 2a. After deproteination, each RNA sample was adjusted to a concentration of 50 µg ml⁻¹, 30 mM CaCl₂ and translated in Ca²⁺-depleted SHE cells. The kinetics of IF mRNA activity are given in Fig. 2b. Similar results were obtained when DEAE-dextran was used to facilitate RNA uptake. The appearance of IF mRNA in the human cells after exposure to rI_h-rC_h is very rapid, peaking at 2–3 h post-induction. More significantly, there is an abrupt decline in IF mRNA activity which corresponds to the shutoff of the IF synthesis.

Sehgal *et al.* have shown that interferon shutoff can be prevented when an RNA inhibitor such as actinomycin D or 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole is given to the cells after induction⁴. In these conditions, synthesis of IF can

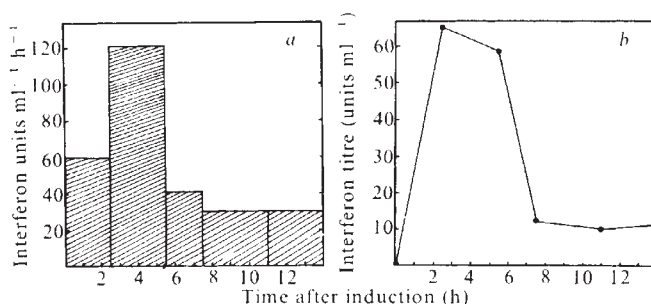


Fig. 2 HF 926 cells in 1/2 gallon roller bottles were induced with 0.4 mM rI_h-rC_h for 1 h then re-fed with D-MEM with 2% foetal calf serum. Media was collected for measuring IF titres and the cells extracted for RNA at 2½, 5½, 7½, 11, and 14 h. Similar results were obtained in two additional experiments. a, Rates of IF synthesis following induction. b, IF mRNA activity following induction.

persist beyond 20 h indicating that the human IF mRNA is a stable species with a half life on the order of 15 h. In view of their finding, it is unlikely that the rapid decline in IF mRNA activity is the result of a natural decay process but is more likely to be the result of a specific inactivation event.

IF mRNA activities during the induction of IF have been examined in two studies by whole cell translational systems. Kronenberg and Friedman¹⁴ measured rI_h-rC_h induced L cell IF mRNA by translation in Vero cells. But, they did not attempt explicitly to correlate the IF mRNA activities with IF protein levels nor did they attempt to study the shutoff process. At 3 h after induction, when no IF was produced, they detected high levels of IF mRNA whereas at 16 h, when IF was synthesised at maximal rate, they observed a reduction in the levels of IF mRNA. These observations indicate that the system of IF induction in mouse cells is different than in human cells and, therefore, complicate comparisons between these two cell types. Furthermore, since no IF mRNA determinations were made during the shutoff of mouse IF, no conclusions can be made concerning the relationship between IF shutoff and IF mRNA activities.

Reynolds and Pitha¹⁵ studied both human IF production and IF mRNA activities using the avian cell translational system originally developed by DeMayer-Guignard *et al.*¹⁶. In contrast to our findings they did not observe a decrease either in the rate of IF production or in the levels of IF mRNA. More recently, Raj and Pitha¹⁷ observed that IF production in human fibroblasts declines 90% by 6–9 h after induction.

An important consideration in the interpretation of these data is the validity of the translational system. An ideal translational system should translate exogenous mRNA with fidelity and high efficiency, thereby reflecting the functional state of the IF mRNA in the originally induced cell. Translation of the human mRNA by another mammalian cell, such as the hamster, should offer the best prospects in this respect.

Work is in progress to identify the regulatory elements involved in the inactivation of IF-mRNA. Such a repression mechanism may be of broader biological relevance.

These results were reported at the 77th annual meeting of the American Society for Microbiology¹⁸. We thank Mr Stanley Lin for many helpful discussions. This research was supported in part by US NIH (grant GM 166066).

Note added in proof: Sehgal *et al.* and Cavalieri *et al.* using the *Xenopus laevis* oocyte translational system have also observed^{19,20} the rapid decline in IF mRNA activity during the shutoff of human IF synthesis.

Table 2 Effect of Ca²⁺-depletion on the translation of human interferon mRNA in SHE cells

Treatment	IF titre (units ml ⁻¹)
Control (media only)	20
Dulbecco's PBS-HEPES (30 min)	19
Ca ²⁺ -free PBS-HEPES (30 min)	76

Translation was as for Table 1 except that before addition of the RNA, the SHE cells were either untreated or incubated with Ca²⁺-free or normal PBS-HEPES buffer for 30 min.

JAMES J. GREENE
CARL W. DIEFFENBACH
PAUL O. P. TS'O

Johns Hopkins University,
Division of Biophysics,
615 North Wolfe Street,
Baltimore, Maryland 21205

Received 22 July; accepted 24 October 1977.

1. Vilček, J. & Havell, E. A. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3909–3913 (1973).
2. Ho, M. & Armstrong, J. A. *Rev. Microbiol.* **29**, 131–161 (1975).
3. Tan, Y. H., Armstrong, J. A., Ke, Y. H. & Ho, M. *Proc. natn. Acad. Sci. U.S.A.* **67**, 464–471 (1970).
4. Sehgal, P. B. & Tamm, I. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1621–1625 (1976).
5. Vilček, J., Havell, E. A. & Kohase, M. *J. infect. Dis.* **133**, A22–A29 (1976).
6. Sehgal, P. B., Tamm, I. & Vilček, J. *Science* **190**, 282–284 (1975).
7. Sehgal, P. B., Tamm, I. & Vilček, J. *Virology* **70**, 532–541 (1976).
8. McCutchan, J. H. & Pagano, J. S. *J. natn. Cancer Inst.* **41**, 351–357 (1968).
9. Dubes, G. R. & Klingler, E. A. *Science* **133**, 99–100 (1961).
10. Graham, F. L. & Van der Eb, A. J. *Virology* **52**, 456–467 (1973).
11. Kohase, M. & Vilček, J. *Virology* **76**, 47–54 (1977).
12. Traub, A., Kaufman, E. & Ginsberg-Tietz, Y. *Expl Cell Res.* **34**, 371 (1964).
13. Penman, S. in *Fundamental Techniques in Virology* (eds. Habel, K. & Salzman, N. P.) 35–48 (Academic, New York, 1969).
14. Kronenberg, L. H. & Friedmann, T. J. *gen. Virol.* **27**, 225–238 (1975).
15. Reynolds, F. H. & Pitha, P. M. *Biochem. Biophys. Res. Commun.* **59**, 1023–1030 (1974).
16. DeMaeyer-Guignard, J., DeMaeyer, E. & Montagnier, L. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1203 (1972).
17. Raj, N. B. K., Pitha, P. M. *Proc. natn. Acad. Sci. U.S.A.* **74**, 1483–1487 (1977).
18. Greene, J. J., Dieffenbach, C. W. & Ts'o, P. O. P. *Annual Meeting of the American Society for Microbiology* Abstract no. S-64 (1977).
19. Sehgal, P. B. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 3409–3413 (1977).
20. Cavaliere, R. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **74**, 4415–4419 (1977).

Bleomycin causes release of nucleosomes from chromatin and chromosomes

THE antibiotic bleomycin is effective against certain human neoplastic diseases¹. Interaction of purified DNA with bleomycin results in the release of free bases from the DNA backbone^{2,3}, disruption of phosphodiester linkage, and destruction of the deoxyribose moiety⁴. Bleomycin-induced DNA strand scissions, both single- and double-strand, were found in bacterial as well as in mammalian systems^{5,6}. But, all previous work on bleomycin has dealt with DNA molecules extracted from treated cells, and analysed on sucrose density gradients. This is not sensitive enough to probe the action of bleomycin with respect to chromatin structure. It is now known that chromatin is composed of a number of repeating subunits connected by a continuing DNA duplex analogous to 'beads on a string'⁷. The bead-like structure, termed nucleosome or nu-body, contains 140 base pairs and eight histone molecules^{8,9}. The DNA between nucleosomes, about 40–60 base pairs in length, is referred to as the linker DNA. Both DNAs in the nucleosomes and in the linker are susceptible to nuclease digestion, but the latter is more sensitive^{10–12}. We have analysed the effects of bleomycin on chromatin with special reference to the characterisation of the breakage sites in a number of mammalian cell systems. The results reported here indicate that bleomycin is one of the few non-enzymatic chemicals which can be used to study the chromatin structure, and that the action of bleomycin is not identical to that of nucleases—bleomycin does not digest nucleosome DNA but attacks only the linker DNA.

When nuclei isolated from cultured Chinese hamster cells (line CHO) were incubated with various concentrations of bleomycin, followed by deproteinisation and gel electrophoresis on agarose, a characteristic series of DNA bands was obtained (Fig. 1). DNA isolated from untreated nuclei showed a high molecular weight (gel 3), indicating that endogenous nucleases had no significant effect on native DNA. The DNA isolated from the nuclei treated with higher concentrations of bleomycin contained more DNA fragments with lower molecular weight than that from

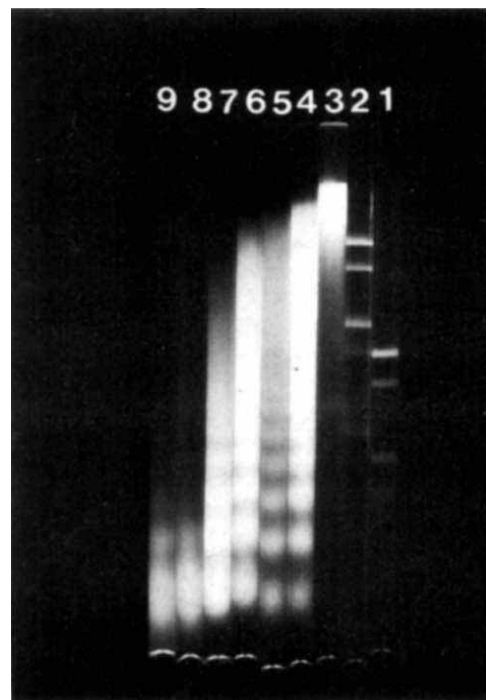


Fig. 1 Agarose gel analysis of DNA digested *in situ* with bleomycin. Approximately 10^7 CHO nuclei isolated using the method described previously²⁰ were incubated in 400 μ l of reaction mixture containing 15 mM Tris-HCl, pH 8.0, 0.34 M sucrose, 0.5 mM EDTA and 15 mM 2-mercaptoethanol and various concentrations of bleomycin at 22 °C for 2 h. The reaction was terminated by the addition of an equal volume of solution containing 0.15 M NaCl, 15 mM EDTA, 50 mM Tris-HCl, pH 8.25, 0.3% sarcosyl NL 97. The mixture was treated with 100 μ g ml⁻¹ of heated (100 °C for 5 min) RNase A at 37 °C for 1 h. DNA was extracted by deproteinisation with phenol saturated with 0.15 M NaCl, 0.1 M EDTA followed by chloroform:isoamylalcohol (24:1, v/v), and precipitated with 2.5 volumes 100% ethanol. Agarose (1.4%) was made in E buffer containing 40 mM Tris-HCl, pH 8.0, 5 mM Na-acetate, 1 mM EDTA. Electrophoresis was run in E buffer at 4 mA per gel. Gel (1) *Hae*III-digested PM 2 DNA, (2) *Hind*III-digested PM 2 DNA, (3–9) CHO DNA; (3) control, no bleomycin added, (4) 1.5 μ g/ml, (5) 15 μ g ml⁻¹, (6) 300 μ g ml⁻¹, (7) 1 mg ml⁻¹, (8) 2 mg ml⁻¹, and (9) 3 mg ml⁻¹ bleomycin.

nuclei treated with lower concentrations of bleomycin, indicating that the fragmentation of nuclear DNA is dose dependent (gels 4–9).

The molecular weight of each band (Fig. 1) was determined by three methods. (1) Construction of a calibration curve by plotting the migration distance of the PM 2 phage DNA fragments digested by restriction enzymes *Hind*III and *Hae*III (Fig. 1, gel 1, 2) against their molecular weights (data not shown). The lengths of the second band (from the bottom of the gels) to the 5th band, averaged from three independent experiments, were 343 ± 7 , 535 ± 15 , 693 ± 15 , and 890 ± 10 base pairs, respectively. These values correspond well to a monomer unit length of 175 ± 4 base pairs. (2) Linear regression analysis of the logarithms of the molecular weight of the restriction fragments plotted against mobilities. The repeated unit length was found to be 170 ± 8 base pairs. (3) Subtracting the length of each band from that of the next higher band. Using this method, the unit length was found to be 182 ± 10 base pairs. The average of these three determinations yielded a unit length of 175 ± 8 base pairs, which is in good agreement with the nucleosome unit length reported by other investigators^{13,14} using nuclease digestions. Bleomycin, therefore, acts on chromatin in a manner analogous to that of nucleases.

Further analyses showed, however, that the modes of action on chromatin DNA between nucleases and bleomycin are not exactly the same. Nucleases have been shown

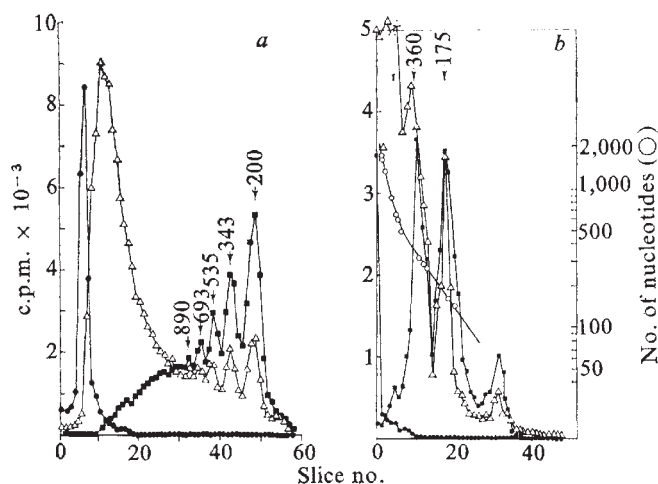


Fig. 2 Gel electrophoretic patterns of DNA isolated from bleomycin-treated CHO nuclei. CHO cells were labelled with ^3H -TdR ($0.5 \mu\text{Ci ml}^{-1}$, 22 Ci mmol^{-1}) for 24 h. Procedures for isolation of nuclei and digestion by bleomycin are described in Fig. 1. ●, Control, no bleomycin; △, $1.5 \mu\text{g ml}^{-1}$ bleomycin for 5 h; ■, $300 \mu\text{g ml}^{-1}$ for 5 h. *a*, DNA samples were electrophoresed in non-denaturing conditions as described in Fig. 1. *b*, Corresponding samples as run in denaturing conditions, that is, 5% acrylamide in 99% formamide²¹. The marker DNAs (○) shown in (*b*) are *Hind*III- and *Hae*III-digested PM 2 DNA fragments^{22,23}. The arrows represent the corresponding numbers of base pairs (*a*) or nucleotides (*b*) as calibrated from the same DNA markers co-electrophoresed.

to digest intra-nucleosomal DNA and produce a series of bands with a multiplicity of 10 nucleotides as analysed by a denaturing gel¹⁵⁻¹⁷. We did not obtain similar results when bleomycin was used.

Figure 2 shows gel electrophoretic patterns of radioactively-labelled DNA isolated from bleomycin-treated nuclei or control nuclei. The specific activity of DNA was approximately $20,000 \text{ c.p.m. } \mu\text{g}^{-1}$, sensitive enough to detect an amount as small as 5 ng. The control sample gave no detectable amount of DNA with molecular weight less than 5,000 base pairs. The treated sample showed DNA in a series of bands with multiplicity of 175 base pairs as described previously.

When the same samples shown in Fig. 2*a* were denatured and separated on denaturing gels, a characteristic pattern emerged (Fig. 2*b*). In the bleomycin-treated samples, only bands with sizes of 50, 175 and 360 nucleotides were present. The minor band (50 nucleotides in length as analysed by extrapolation of the calibration curve) was not present in the undenaturing DNA gel. This fragment must, therefore, have come from a higher molecular weight, double-stranded duplex containing single-strand nicks.

That the double-stranded, high molecular weight DNA contains single-strand nicks can also be seen by comparing the same sample ($300 \mu\text{g ml}^{-1}$, 5 h) electrophoresed in non-denaturing (Fig. 2*a*) and denaturing conditions (Fig. 2*b*). Figure 2*a* shows that about 50% of the counts migrated with molecular weights higher than 500 base pairs. But, when this sample was denatured and separated in denaturing gel, the high molecular weight DNA (>500 base pairs) disappeared and was replaced by monomeric and dimeric fragments. From these results, we conclude that the high molecular weight DNA duplex contained single-strand breaks spaced regularly, probably every 50, 175 or 360 nucleotides apart.

We detected no multiplicity of 10 nucleotides on the gels. This indicates that the nuclease-sensitive sites within nucleosomes are not sensitive to bleomycin. An explanation for this phenomenon is that bleomycin is more specific

in its recognition of sequence or structure of the DNA helix than are nucleases. Previously, we have reported that DNA with molecular weight less than approximately 12 nucleotides is resistant to bleomycin degradation, regardless of the sequence¹⁸. If DNA within the nucleosomes is kinked every 10 base pairs as proposed by Crick and Klug¹⁹, or embedded within protein molecules, then the intranucleosomal DNA may become resistant to bleomycin activity, so that no internal nicks can be created. Our interpretation is that bleomycin induces damage only in the internucleosomal linker DNA, and that the 50 nucleotide pieces probably came from nicks in the linker DNA at or near the adjacent nucleosomes. We also found that several mammalian cells in addition to CHO responded to bleomycin in a similar manner, and that the antineoplastic protein, neocarzinostatin, also yielded DNA degradation products similar to those of bleomycin. We believe that these drugs will be very useful tools in the study of chromatin structure.

This work was supported in part by the National Institute of Environmental Health Science (grant ES-01304) and the US NSF (ENV-76-82241). We thank Drs Roger Hewitt, R. T. Kovacic and James C. Wang for supplying PM 2 bacteriophage DNA.

M. TIEN KUO

T. C. HSU

Department of Biology,
The University of Texas System Cancer Center,
M.D. Anderson Hospital and Tumor Institute,
Houston, Texas 77030

Received 6 September; accepted 24 October 1974.

- Blum, R. H., Carter, S. K. & Agre, K. *Cancer* **31**, 903-914 (1977).
- Haidle, C. W., Weiss, K. K. & Kuo, M. T. *Molec. Pharmacol.* **8**, 531-537 (1972).
- Muller, W. E. G., Yamazaki, Z.-I., Breter, H. J. & Zahn, R. K. *Eur. J. Biochem.* **31**, 518-525 (1972).
- Kuo, M. T. & Haidle, C. W. *Biochim. biophys. Acta* **335**, 109-114 (1973).
- Suzuki, H., Nagai, K., Akutzu, E., Yamaki, H., Tanaka, N. & Umezawa, H. *J. Antibiot.* **23**, 473-480 (1970).
- Terasima, T., Yasukawa, M. & Umezawa, H. *Gann* **61**, 513-516 (1970).
- Olins, A. L. & Olins, D. E. *Science* **183**, 330-332 (1974).
- Kornberg, R. D. *Science* **184**, 868-871 (1974).
- Simpson, R. T. & Bustin, M. *Biochemistry* **15**, 4305-4312 (1976).
- Sollner-Webb, B. & Felsenfeld, G. *Biochemistry* **14**, 2915-2920 (1975).
- Shaw, B. R., Herman, T. M., Kovacic, R. T., Beaudreau, G. S. & Van Holde, K. E. *Proc. natn. Acad. Sci. U.S.A.* **73**, 505-509 (1976).
- Axel, R. *Biochemistry* **14**, 2921-2925 (1975).
- Rill, R. L., Nelson, D. A., Oosterhof, D. K. & Hozier, J. C. *Nucleic Acid Res.* **4**, 771-789 (1977).
- Compton, J. L., Ballard, M. & Chambon, P. *Proc. natn. Acad. Sci. U.S.A.* **73**, 4382-4386 (1976).
- Noll, M. *Nucleic Acid Res.* **1**, 1573-1578 (1974).
- Whitlock, J. P., Jr, Rushizky, G. W. & Simpson, R. T. *J. biol. Chem.* **252**, 3003-3006 (1976).
- Sollner-Webb, B. & Felsenfeld, G. *Cell* **10**, 537-547 (1977).
- Kuo, M. T., Haidle, C. W. & Inners, L. D. *Biophys. J.* **13**, 1295-1306 (1973).
- Crick, F. M. & Klug, A. *Nature* **255**, 530-533 (1975).
- Kuo, M. T., Sahasrabudhe, C. G. & Saunders, G. F. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1572-1575 (1976).
- Maniatis, T., Jeffrey, A. & Van de Sande, H. *Biochemistry* **14**, 3787-3794 (1975).
- Brack, C., Eberle, M., Bieckle, T. A. & Yuan, R. *J. molec. Biol.* **104**, 305-309 (1976).
- Kovacic, R. T. & Van Holde, K. E. *Biochemistry* **16**, 1490-1498 (1977).

Dissociation of 'hybrid' isozymes on electrophoresis

HETEROMERIC or 'hybrid' isozymes can be readily demonstrated by electrophoretic techniques. They can be observed, for example, in individuals who are heterozygous for alternative alleles at an autosomal locus determining a multimeric enzyme or where multiple loci are involved in the determination of a multimeric enzyme. Heteromeric forms can also be generated by techniques such as somatic cell fusion, and by dissociation and recombination *in vitro*. The number of heteromeric isozymes displayed by a particular enzyme is related to its subunit composition. For example, in a heterozygote a dimeric enzyme exhibits one heteromeric isozyme and two homomeric isozymes, whereas a

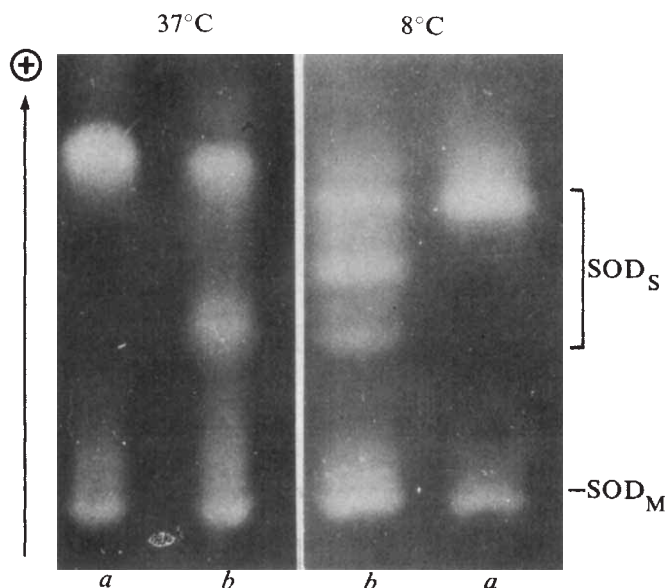


Fig. 1 Starch gel electrophoresis of SOD at 8 °C and at 37 °C using a Tris, EDTA, maleate, MgCl_2 buffer system at pH 7.4. Electrophoresis was carried out for 17 h at 5 V cm^{-1} at 8 °C and 3.5 V cm^{-1} at 37 °C. Temperature was regulated during electrophoresis by circulating water at the appropriate temperature through electrically insulated metal plates which were in close contact with the gel surface. The gels were stained using MTT 0.1 mg ml^{-1} and phenazine methosulphate 0.05 mg ml^{-1} in Tris/Cl buffer, pH 8.0 applied as an agar overlay. The samples are liver extracts from (a) an individual of phenotype $\text{SOD}_s 1$ and (b) an individual of phenotype $\text{SOD}_s 2-1$. The most cathodal zone of activity is SOD_M which does not show the same genetic variation.

tetrameric enzyme displays three heteromeric isozymes in addition to the two homomeric forms. Thus, the number of heteromeric isozymes provides a valuable method of assessing the number of subunits making up active enzyme molecules and indeed this approach has recently been used in a survey of the subunit structures of 66 different human enzymes¹. It was found that heteromeric isozymes usually occur as described here, but there are occasional exceptions where the expected 'hybrid' isozymes are unusually weak or absent. We have investigated one component of the electrophoretic separation system—that is, temperature—and its influence on the occurrence of heteromeric isozymes using soluble superoxide dismutase (SOD_s) as a test case. This was prompted by a report that this dimeric enzyme is

readily dissociated at high temperatures²; also, the heteromeric isozymes of SOD_s can be conveniently studied by electrophoretic methods. Our results suggest that temperature has a crucial effect on the occurrence of the heteromeric isozymes of this enzyme.

SOD , which catalyses the disproportionation of superoxide radicals, exists in two forms in eukaryotes³. One form, SOD_s , found in the cytosol, is a dimeric copper-zinc protein with a molecular weight of 32,000. The mitochondrial form, SOD_M , is a tetrameric manganese protein with a molecular weight of about 72,000. An electrophoretically slow genetic variant of the soluble enzyme, designated $\text{SOD}_s 2-1$ and attributable to heterozygosity for a variant allele SOD_s (ref. 2), and the common allele SOD_s^1 , has been identified⁴. In heterozygotes, three isozymes are observed under standard electrophoretic conditions. The two outer zones representing the homomeric forms, and the middle zone the heteromeric ('hybrid') form containing polypeptides determined by each allele.

A liver extract from an individual of the $\text{SOD}_s 2-1$ phenotype was subjected to electrophoresis in starch gels at various temperatures ranging from 8–55 °C. At 8 °C, the usual three-banded pattern was observed. At 20–30 °C, a three-banded pattern was also observed; but, whereas the homomeric isozymes were of usual activity, the heteromeric form was much less active. At temperatures above 35 °C, two bands were observed with mobilities corresponding to the two homomeric forms and displaying normal activity. The heteromeric isozyme was not present (Fig. 1). Similar results were obtained with man mouse hybrid material derived from somatic cell fusion and from *in vitro* dissociation and recombination (Fig. 2).

A simple explanation for these findings would be that heteromeric isozymes of SOD_s are more heat sensitive than homomeric forms. Heating the SOD_s isozymes (40 °C, 17 h) in the gels using the technique of McAlpine *et al.*⁵, subsequent to their separation by electrophoresis at 8 °C, did not, however, lead to any noticeable loss of activity. Furthermore, heating a crude tissue extract at 70 °C for up to 3 h, followed by electrophoresis at 8 °C, revealed that the homomeric and heteromeric isozymes lose activity at the same rate. Thus, it seems that the disappearance of the heteromeric isozymes during electrophoresis is not due to heat lability.

A more likely explanation (similar to that proposed to account for the disappearance on electrophoresis of the asymmetric 'hybrid' forms of haemoglobin^{6,7}) is that the dimeric SOD_s isozymes are in rapid equilibrium with their constituent polypeptides at elevated temperatures:

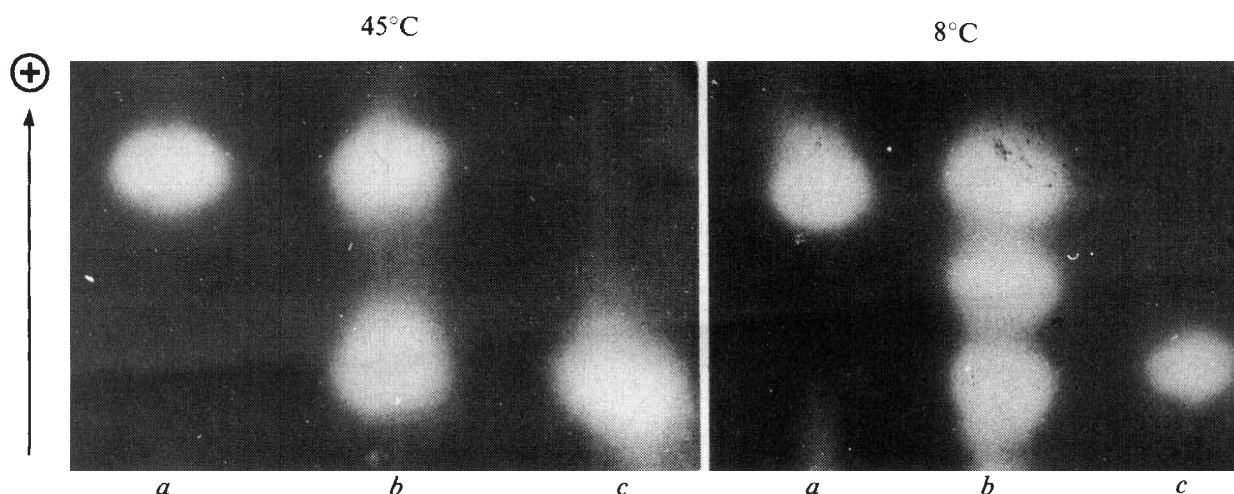
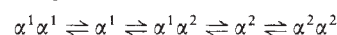


Fig. 2 Electrophoresis of SOD_s at 8 °C (5 V cm^{-1}) and at 45 °C (3 V cm^{-1}) using the procedure detailed in Fig. 1. Samples: a, human liver; b, a 'hybrid' preparation of man and mouse SOD_s prepared by *in vitro* dissociation and recombination; c, mouse liver.

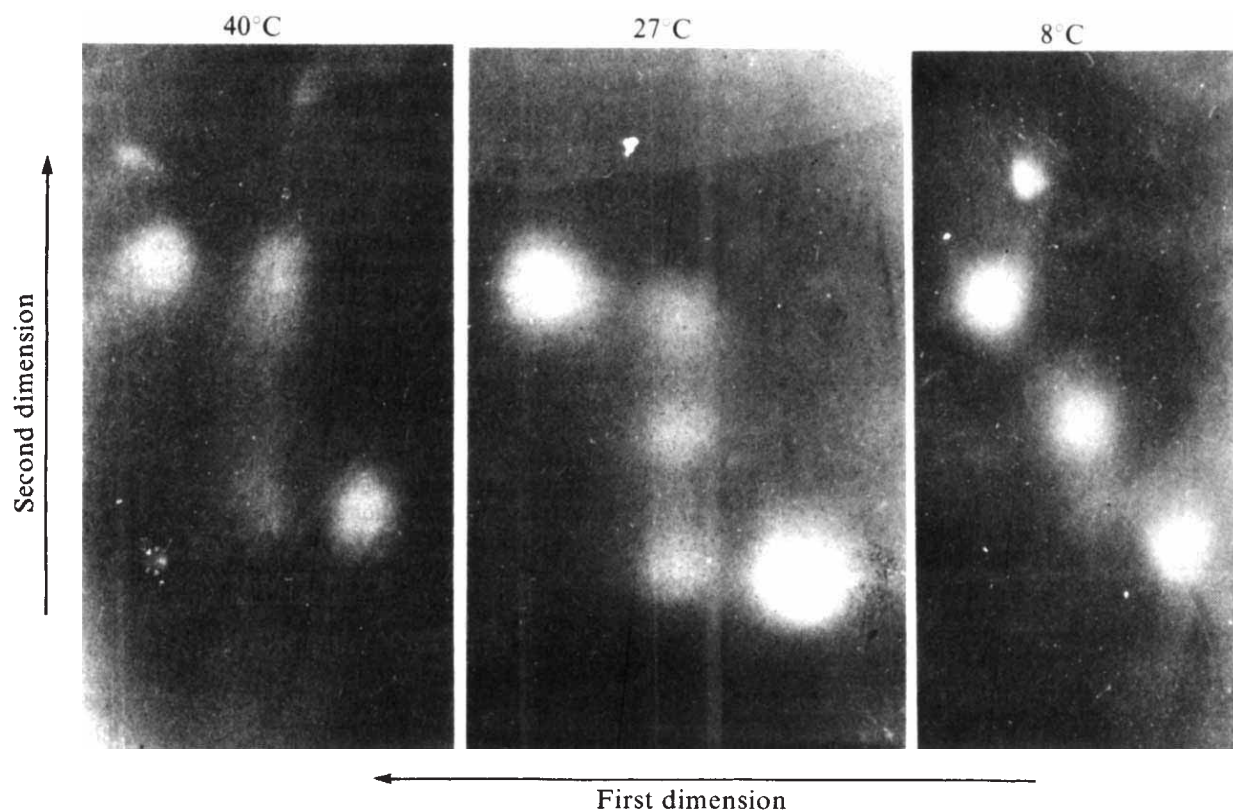
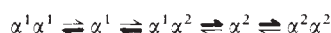


Fig. 3 Two-dimensional electrophoresis of man-mouse 'hybrid' SOD₅ isozymes. In each case electrophoresis in the first dimension was at 8 °C. In the second phase electrophoresis was carried out at 8, 27 and 40 °C. Conditions for electrophoresis and SOD detection as described in Fig. 1.

On electrophoresis, this equilibrium is disturbed due to the separation of the unlike forms, so that there is a progressive depletion of the heteromeric ('hybrid') isozyme:



In ordinary temperature conditions, the dissociation-reassociation process is presumably much less rapid, and the heteromeric SOD₅ isozyme persists.

This interpretation has been investigated using a two-dimensional technique. In the first dimension, separation was carried out at 8 °C (that is, the usual condition for SOD₅ electrophoresis) and then in the second dimension at 8 °C, 45 °C or some intermediate temperature. In one series of experiments, samples containing three isozymes of SOD₅ were separated by electrophoresis at 8 °C and then the pieces of gel containing the middle band (the heterodimeric isozyme) were cut out and subjected to electrophoresis at 45 °C. In each case—regardless of whether the heterodimer was derived from a SOD₅ 2-1 heterozygote, a somatic cell hybrid or from *in vitro* dissociation and recombination of man and mouse SOD₅—this single band was found to generate two isozymes of equal intensity and with electrophoretic mobilities identical to the corresponding homodimers. In another series of experiments, the second phase of electrophoresis was carried out using all three isozymes (the two homodimers and the intermediate heterodimer). With increasing temperature in the second phase of electrophoresis, there was a coincident apparent increase in the rate at which the heteromeric isozyme generated the homomeric bands (Fig. 3). Thus, it seems that temperature has a crucial influence on the appearance of the heteromeric isozymes of SOD₅ after electrophoresis.

We have made preliminary investigations into the effects of temperature on a limited number of other enzymes which exhibit heteromeric isozymes, but so far no example of dissociation during electrophoresis which is as clear-cut as that described for SOD₅, has been seen. A wide range of enzymes, however, remains

to be investigated as well as many other factors, such as pH, buffer ion concentration and coenzyme binding, which might affect the dissociation of isozymes during electrophoresis. Also, the condition of the sample may influence the appearance of heteromeric forms, since it has recently been shown that, although an asymmetric 'hybrid' haemoglobin band is not usually seen—for example, in A-S and A-C heterozygotes—on conventional electrophoresis, this type of band is observed when the haemoglobins are reduced and maintained in anaerobic conditions during electrophoresis⁸.

It seems particularly worthwhile to investigate any multimeric enzymes (or other proteins) which are found to show exceptional isozyme patterns in respect of 'hybrid' forms in case the hybrids are being lost by dissociation during the separation process. For instance, in the case of galactokinase there is evidence for a dimeric structure⁹, but a heteromeric isozyme with a mobility intermediate between the human and mouse isozymes has not been observed in hybrid cell lines¹⁰. Another example is enolase, a dimeric enzyme determined at three separate loci, coding for the polypeptides α , β and γ . The heterodimers $\alpha\beta$ and $\gamma\beta$ are observed after electrophoresis together with the homodimers $\alpha\alpha$, $\beta\beta$ and $\gamma\gamma$, but the expected heterodimer $\alpha\gamma$ has not been detected in either human or rat tissues¹¹.

In every unusual case, however, it will be necessary to bear in mind other reasons for the absence of heteromeric isozymes after electrophoresis. In the case of enzymes determined by X-linked loci, where only one X chromosome is functioning in each cell, there is no opportunity for unlike polypeptides to come together to form hybrid isozymes. This is illustrated by the glucose-6-phosphate dehydrogenase isozymes in females of the AB phenotype: hybrid isozymes have been observed in oocytes from such females, in which both X chromosomes are active, but not in other tissues¹². Differences in gene expression from one cell type to another also results in the absence of heteromeric forms. For example, lactate dehydrogenase (LDH) AC and BC hybrids are not found in testicular tissue. This is because only the LDH_C locus

is active in primary spermatocytes, whereas the LDH_A and LDH_B loci are expressed in other cells of the testis¹³.

Sometimes, there may be stereochemical reasons why dissimilar polypeptides do not form a hybrid. A possible example is the rare genetic variant of human LDH¹⁴. The multimeric enzymes which exhibit mitochondrial and cytoplasmic forms, determined by separate loci, may provide other examples of stereochemical constraints. Presumably, the gross structural difference between SOD_S and SOD_M—one a dimer, the other a tetramer—prevents the formation of a cytoplasmic-mitochondrial hybrid form. But the reasons why other multimeric enzymes with a similar subcellular distribution (for example, malate dehydrogenase, aspartate aminotransferase, malic enzyme) do not exhibit cytoplasmic-mitochondrial hybrid isozymes, are obscure. In most cases, these isozymes have the same subunit structure and in at least one case (aspartate aminotransferase), a fair degree (about 48%) of homology in amino acid sequence^{15,16}. Perhaps the hybrid forms occur *in vivo* but are dissociated *in vitro*.

We thank Anthony Baker for technical help.

Y. H. EDWARDS
D. A. HOPKINSON
HARRY HARRIS*

MRC Human Biochemical Genetics Unit,
The Galton Laboratory,
University College,
Gower Street, London WC1, UK

Received 19 July; accepted 3 November 1977.

*Present address: Department of Human Genetics, University of Pennsylvania.

- Hopkinson, D. A., Edwards, Y. H. & Harris, H. *Ann. hum. Genet.* **39**, 4383–4411 (1976).
- Tegelström, H., *Hereditas* **81**, 185–197 (1975).
- Fridovich, I. A. *Rev. Biochem.* **44**, 147–159 (1975).
- Beckman, G. & Pakarinen, A. *Hum. Hered.* **23**, 346–351 (1973).
- McAlpine, P., Hopkinson, D. A. & Harris, H. *Ann. hum. Genet.* **34**, 61–71 (1970).
- Guidotti, G., Königsberg, W. & Craig, L. C. *Proc. natn. Acad. Sci. U.S.A.* **50**, 774–782 (1963).
- Benesch, R., Benesch, R. E. & Tyuma, I. *Proc. natn. Acad. Sci. U.S.A.* **56**, 1268–1274 (1966).
- Bernstein, S. C. & Bowman, J. E. *Biochim. biophys. Acta* **427**, 512–519 (1976).
- Blume, K. G. & Beutler, E. *J. biol. Chem.* **246**, 6507–6510 (1973).
- Elsevier, S. M. *et al. Nature* **251**, 633–635 (1974).
- Pearce, J., Edwards, Y. H. & Harris, H. *Ann. hum. Genet.* **39**, 263–276 (1976).
- Gartler, S. M., Liskay, R. M., Campbell, B. K., Sparkes, R. & Gant, N. *Cell Differentiat.* **1**, 215–216 (1972).
- Blanco, A. & Zinkham, W. H. *Science* **139**, 601–602 (1963); Goldberg, E. *Science* **139**, 602–603 (1963).
- Nance, W. E., Claffin, A. & Smithies, O. *Science* **142**, 1075–1077 (1963).
- Kagamiyama, H., Sakakibara, R., Wada, H., Tanase, S. & Morino, Y. *J. Biochem.* **82**, 291–294 (1977).
- Barra, D. *et al. FEBS Lett.* (in the press).

N-methylphenylalanine at the N-terminus of pilin isolated from *Pseudomonas aeruginosa* K.

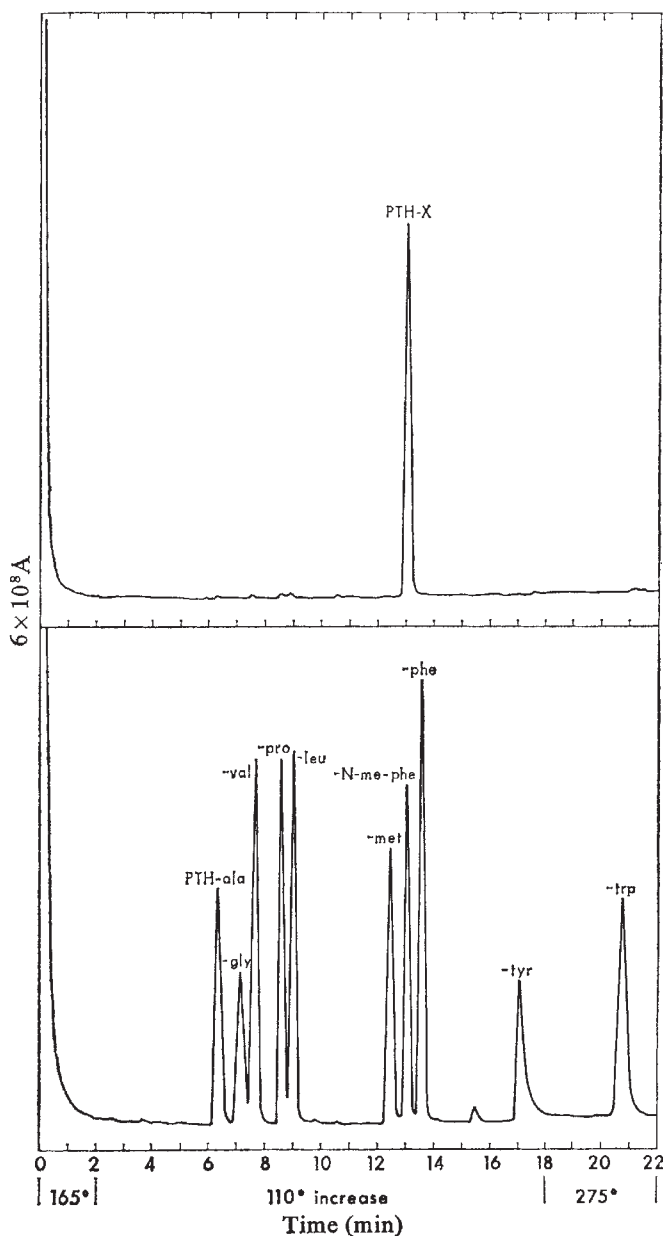
AUTOMATED amino acid sequence analysis of pili protein isolated from the multi-piliated 2Pfs mutant of *Pseudomonas aeruginosa* strain K (PAK)¹ has shown that N-methylphenylalanine is the amino terminus. These pili, which mediate the infectious process of several *Pseudomonas* bacteriophages², are composed of one protein subunit, pilin, of molecular weight 17,800 (ref. 1). Mass spectroscopy (MS), 270 MHz proton magnetic resonance spectroscopy (PMR), ultraviolet absorption spectroscopy, gas/liquid chromatography (GLC) and thin layer chromatography (TLC) have demonstrated the identity of the phenylthiohydantoin (PTH) derivative of the amino terminal amino acid (PTH-X) to be identical with synthetic PTH-N-methylphenylalanine.

N-methylphenylalanine was synthesised by methylation of tertiary butyloxycarbonyl-phenylalanine dicyclohexylamine salt with methyl iodide in tetrahydrofuran in the presence of sodium hydride for 24 h at 21 °C (ref. 3), with subsequent removal of the carboxyl- and amino-protecting groups by treatment with aqueous citric acid, pH 3, and 50% (v/v) trifluoroacetic acid/dichloromethane, respectively. Reaction of N-methyl-

phenylalanine (90 mg) with 5% (v/v) phenylisothiocyanate in pyridine (2.4 ml) at 55 °C for 1 h gave the phenylthiocarbonyl derivative, which after drying in a vacuum, was converted to the PTH derivative by reaction with 1 M HCl (1 ml) at 80 °C for 10 min, the crude product being obtained by extraction with ethyl acetate (3 × 1 ml). Recrystallisation from diethyl ether afforded colourless crystals m.p. 113 °C.

Since the PTH-N-methylphenylalanine synthesised coeluted as a single peak on GLC with PTH-X, the product of the first Edman degradation of PAK/2Pfs pilin (Fig. 1), and cochromatographed as a single spot on silica gel TLC (E. Merck, precoated) using two solvent systems (*R_f* 0.94 with 1,2-ethanedichloride:acetic acid 30:7 (v/v)⁴ and *R_f* 0.33 with chloroform:ethanol, 98:2 (v/v)⁵), MS, PMR and ultraviolet absorption spectroscopy were carried out to compare the two compounds further. To obtain sufficient PTH-X for PMR analysis

Fig. 1 GLC of (a) PTH-X (2% of the product from the first Edman degradation of 5 mg (280 nmol) pilin, and (b) PTH-N-methylphenylalanine, together with a number of standard PTH-amino acids (5 nmol of each). Separations were made on a Beckman GC-45 gas chromatograph fitted with a 1.2 m × 2 mm internal diameter column of 10% SP-400 on 100–120 mesh Supelcoport (Supelco), using helium carrier gas at 40 ml min⁻¹.



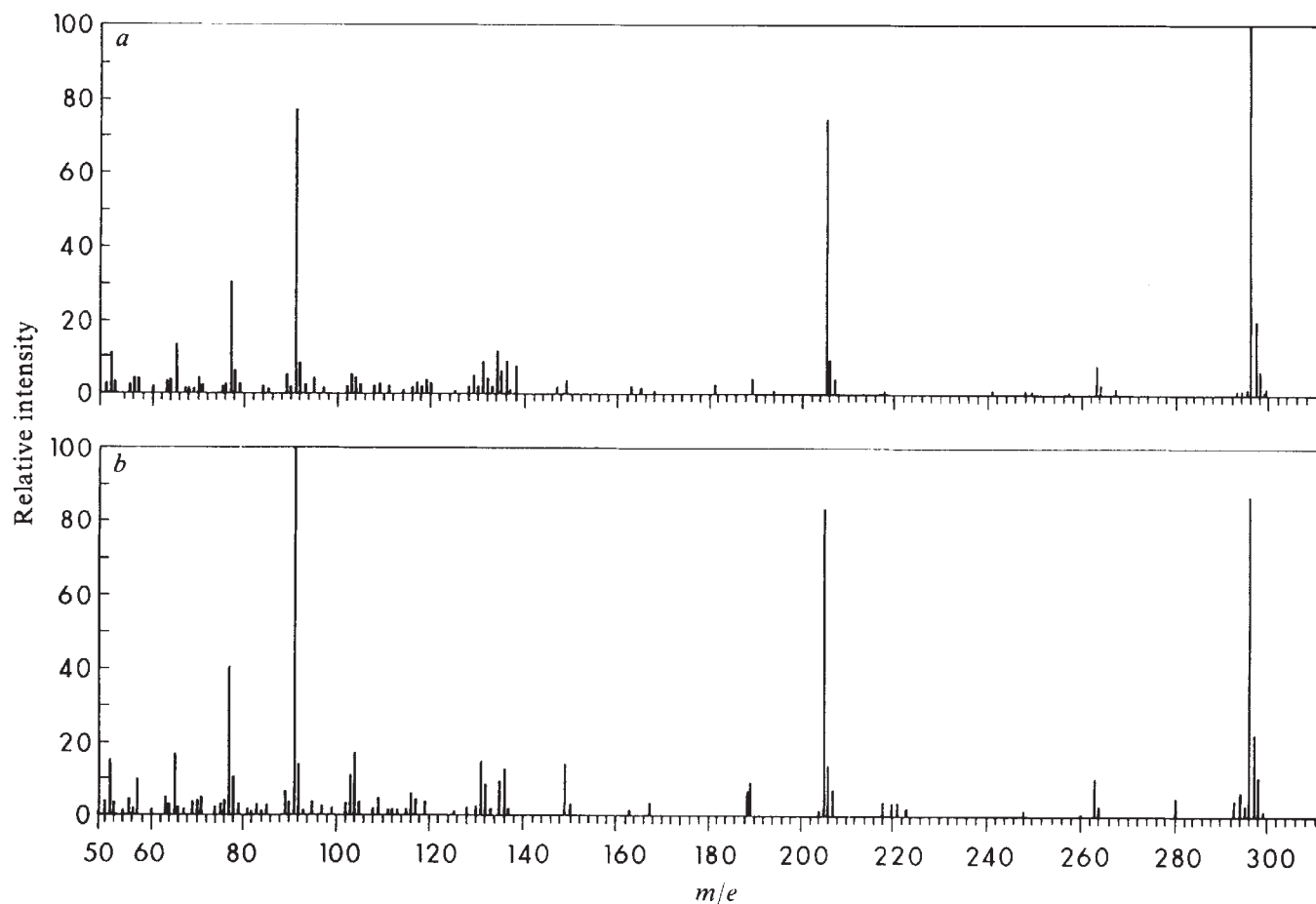
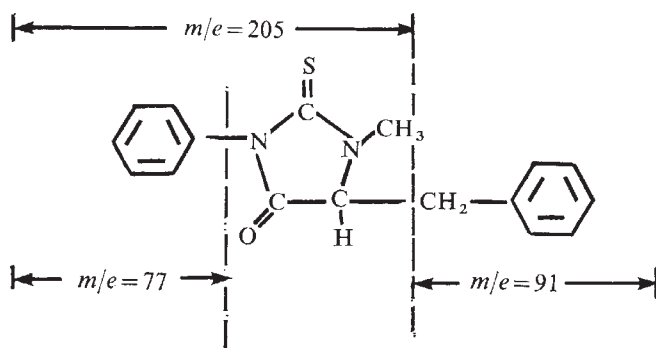


Fig. 2 MS of PTH-*N*-methylphenylalanine (a) and PTH X (b). Both spectra were obtained on 10 μ g or less of sample by direct probe, 70 eV electron impact ionisation at an initial ion source temperature of 150 °C, using an AEI MS50 high resolution mass spectrometer in conjunction with a DS50 computer.

(0.2 mg), a combined total of 35 mg (2 μ mol) pilin were subjected to one Edman degradative cycle.

The MS of the two compounds is shown in Fig. 2. A molecular ion of $m/e = 296$ and major fragments of $m/e = 205$, 91 and 77 were evident in both spectra.



Similarly, Fig. 3 provides a comparison of the compounds' PMR spectra. The peaks downfield from 6.7 to 7.4 p.p.m. represent the two phenyl rings in the PTH derivative. The peak at 4.45 p.p.m. is due to the α -carbon proton, whereas the peak at 3.33 p.p.m. is due to the β -carbon protons. The singlet at 3.42 p.p.m. is indicative of the *N*-methyl protons.

Ultraviolet absorption spectra of the compounds showed that both have a high $\epsilon_{245}/\epsilon_{269}$ ratio of 0.64 compared with values at about 0.4 for PTH derivatives of other amino acids, with the exceptions of 0.72 for PTH-*N*^ε-phenylthiocarbamyl-lysine and

0.67 for PTH-proline⁴. In addition, the ϵ_{269} molar absorption coefficient for the synthetic compound is lower at 13,500 than the usual value of around 16,000 for these amino acid derivatives⁴.

Attempts to regenerate and detect by amino acid analysis *N*-methylphenylalanine by hydrolysis of PTH X and the synthetic derivative were unsuccessful. Neither acidic hydrolysis with 5.7 M HCl at 110 °C for 24 h or 57% HI at 127 °C for 20 h, nor alkaline hydrolysis with 0.2 M NaOH/0.1 M Na₂S₂O₄ at 127 °C for 4 h (ref. 6) yielded *N*-methylphenylalanine. To be able to detect this amino acid the method of Coggins and Benoiton⁷ in which a halved buffer and ninhydrin flow through the reaction coil, resulting in a double reaction time with ninhydrin, was used.

Whereas *N*-methylphenylalanine is a novel component of protein, it has been shown to be a constituent of the peptide antibiotic Staphylomycin S, produced by *Streptomyces virginiae*⁸. Pettigrew and Smith⁹ have reported the finding of dimethylproline as a blocking group at the amino terminus of *Crithidia oncopelti* cytochrome C557 and speculated on the possibility of other α -*N*-methylated amino acids existing in proteins. In the case of PAK/2PIS pilin, however, the *N*-methylphenylalanine did not act as a blocking group to sequence analysis but merely presented a new PTH amino acid, which was not readily identified. Edman degradation in subsequent cycles proceeded normally with no obvious drop in efficiency between the first and second cycles. This is not surprising, since proline, the only other secondary amine usually encountered in protein sequence work, is not known to interfere with the process. (These sequence data are beyond the scope of this communication and will be presented separately.)

Finally, although conventional amino acid analysis of the 24-h 5.7 M HCl protein hydrolysate failed to detect the new amino

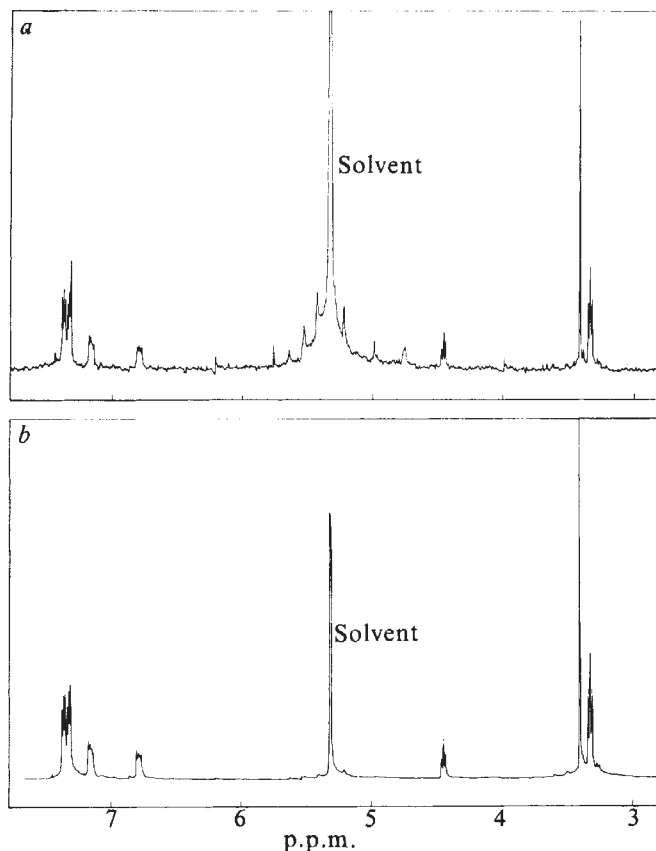


Fig. 3 PMR spectra of PTH-X (a) and PTH-N-methylphenylalanine (b). Both samples, 0.2 and 1.0 mg respectively, were washed with 99.8% D_2O and dried *in vacuo* (twice) before being dissolved in 0.4 ml $CD_2Cl_2-d_2$. The spectra were recorded with a Bruker HX 270 superconducting NMR spectrometer at 270 MHz and ambient temperature. Comparative peaks were adjusted to similar height. Chemical shifts are reported in p.p.m. downfield from tetramethylsilane.

acid¹, application of the method of Coggins and Benoiton⁷ permitted the detection of 1 mol *N*-methylphenylalanine per 12.6 mol leucine. Thus, on the basis of the previously reported amino acid analysis¹, which established 14 leucine residues per polypeptide (molecular weight 17,800), it is clear there is only one *N*-methylphenylalanine per polypeptide and, as already demonstrated, it occupies the amino terminus of the molecule.

This work was supported by the MRC of Canada. We thank R. S. Hodges, R. U. Lemieux and L. B. Smillie for discussions and for the use of their facilities (L.B.S.). We also acknowledge the technical assistance and expertise of M. Natriss for amino acid analyses, B. D. Sykes for PMR spectra and A. Hogg for mass spectra.

L. S. FROST
M. CARPENTER
W. PARANCHYCH

Department of Biochemistry,
University of Alberta,
Edmonton, Alberta, Canada T6G 2H7

Chaos in the Zhabotinskii reaction

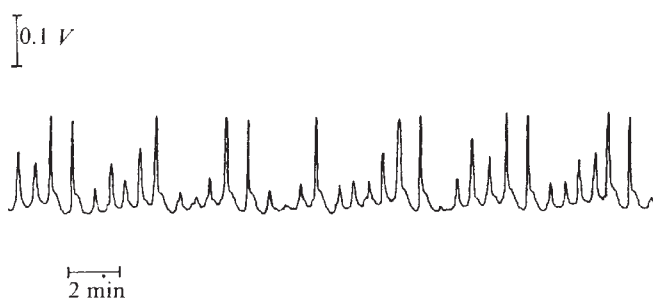
THE Belousov-Zhabotinskii reaction is a chemical Bonhoeffer-van der Pol circuit, that is, a relaxation oscillator that can be run as both an astable and a monostable 'flip-flop'¹⁻³. Apparently the reaction also belongs to the slightly more complicated class of 'universal circuits'⁴ as introduced by Khaikin^{5,6}. Oscillators of this type not only show 'smooth' and 'relaxation type' oscillations^{5,6}, but also 'chaotic' oscillations⁴. As evidenced by the simplest equation of this type⁷, both 'spiral type' and 'screw type' chaos²⁰ are possible in such systems. We present here preliminary evidence for the occurrence of screw-type chaos in the Zhabotinskii reaction.

The system was investigated in isothermal, continuous stirred flow conditions. The initial concentration of reactants was 1.8 mM of $MnSO_4$, 5.6 mM H_2SO_4 , 0.36 M malic acid, and 45 mM potassium bromate. The volume of the reaction mixture was 36 ml. Specifically, the reaction mixture contained 50 parts of a solution 'I' (containing 0.34 g $MnSO_4 \cdot H_2O$, 153 g H_2SO_4 and 54 g malic acid l^{-1}) and 6 parts of a solution 'II' (70 g $KBrO_3 l^{-1}$). The two solutions were continuously injected by means of a peristaltic pump (Pharmacia P-3) at the same (50:6) ratio: total flow rate was 1 ml min^{-1} . The temperature was kept at 30 °C (± 0.01). The volume was kept constant by means of an overflow mechanism. The electrochemical potential was measured by means of a platinum and an Ag/AgCl electrode. The system was allowed 15 min to reach a 'stationary' regime. A subsequent recording over 25 min is presented in Fig. 1. Similar recordings were obtained over several successive hours.

The Zhabotinskii reaction is an (about) 15-variable reaction system⁸⁻¹¹. The electrochemical potential is not uniquely related to one of these variables⁹. Only a limited amount of information can, therefore, be gained from a single measured curve. All that can be said is that the reaction is not periodic over the time interval shown (and over any time interval measured so far). If the observed oscillation corresponds to a limit cycle, the latter is bound to be of 'complicated' type. A complicated limit cycle (being multiply folded in state space) is, however, in itself an indicator for existing chaos. This is because the presence of such a limit cycle in most cases implies the very 'folding' of a cross-section through the trajectorial flow (see refs 4, 7, 20) that is also sufficient for truly non-periodic behaviour¹².

In realistic systems, exogenous perturbations are unavoidable. This means that a limit cycle of large periodicity will not be detectable if its 'region of dominating influence' (where the temporal behaviour is similar to that on the limit cycle) in state space is small. For then very small perturbations (for example, even the small errors of a highly accurate numerical integration routine) suffice to prevent the system from locking into the periodic regime. The longer the period, the greater this sensitivity to perturbations. On the other hand, a very large 'pseudo-random' period (possessing perhaps 10^{10} intermediate maxima,

Fig. 1 Irregular oscillation of electrochemical potential in the Zhabotinskii reaction in isothermal continuous stirred flow conditions. V , electrochemical potential is plotted against t , time.



Received 3 October; accepted 30 October 1977.

1. Frost, L. S. & Paranchych, W. J. *Bact.* **131**, 259-269 (1977).
2. Bradley, D. E. & Pitt, T. L. *J. gen. Virol.* **24**, 1-15 (1974).
3. McDermott, J. R. & Benoiton, N. L. *Can. J. Chem.* **51**, 1915-1919 (1973).
4. Edman, P. in *Protein Sequence Determination* (ed. Needleman, S. B.) 211-255 (Springer, New York, 1970).
5. Bridgen, J., Graffeo, A. P., Karger, B. L. & Waterfield, M. D. in *Instrumentation in Amino Acid Sequence Analysis* (ed. Perham, R. N.) 111-145 (Academic, London, 1975).
6. Smithies, O. et al. *Biochemistry* **10**, 4912-4921 (1971).
7. Coggins, J. R. & Benoiton, N. L. *J. Chromatogr.* **52**, 251-256 (1970).
8. Vanderhaeghe, H. & Parmentier, G. *J. Am. chem. Soc.* **82**, 4414-4422 (1960).
9. Pettigrew, G. W. & Smith, G. M. *Nature* **265**, 661-662 (1977).

which is not unrealistic) is, for all practical purposes, equivalent to a genuinely non-periodic oscillation.

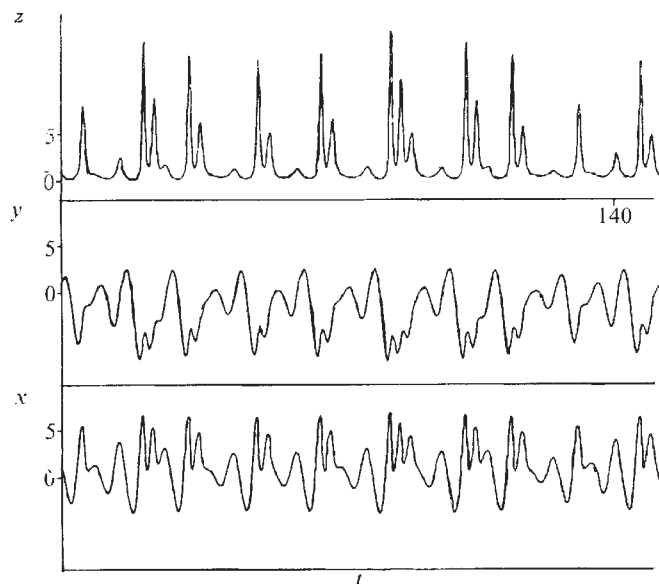
If there is exogenous 'noise' in an apparently chaotic system, the only way to detect an intrinsic chaotic component is through looking at the trajectorial flow in state space. If the flow, although being 'blurred', remains roughly confined to a sub-manifold of detectable shape, and if this shape is one of a few basic types²⁰ known to be chaos-producing in deterministic conditions, then it is possible that the system is also capable of purely 'endogenous' chaos.

Sometimes a single measured variable (or combination of variables, respectively) suffices for this proof. This is the case in spiral-type chaos⁴, for example, if the variable with the largest amplitude is the one that is measured. If subsequent recorded maxima ('amplitudes') are then plotted, each as a function of the respective last, a Li-Yorke-type map is obtained which indicates chaos^{12,13}. Other sub-manifolds in state space (like that determining screw-type chaos^{7,20}) are not so easily detectable, however. The next step in the present investigation will therefore be to simultaneously record several observables, so that three-dimensional stereoscopic pictures of trajectorial flow in observation space can be obtained.

The simplest screw-type chaotic flow⁷ has the temporal behaviour shown in Fig. 2. The time course of the first variable is reminiscent of the flow of Fig. 1. This allows the prediction that typical screw-type chaos can be found in the Zhabotinskii reaction.

The first evidence that the Zhabotinskii reaction is an oscillator of the universal circuit type was provided by the finding of 'double oscillations' in continuous stirred flow conditions^{14,15}. At that time, the dynamical implications of this result were still unknown, however. Later, A. T. Winfree (personal communication) suggested that his observation of a 'meandering' core in a non-stirred excitable version of the Z reagent¹⁶ was evidence for chaos. In the meantime, a similar phenomenon has been found in computer studies of a two-variable excitable medium (ref. 17, Rössler and Kahlert, in preparation); while two variables are sufficient for chaos in non-stirred reaction systems¹⁷⁻¹⁹, three variables are necessary in well-stirred systems. The finding that 'non-planarity' of a limit cycle in three- or higher-dimensional state space (so that a 'folded over' regime is possible) is all that is required for chaos⁴ immediately suggested that virtually all realistic chemical oscillators should be capable of chaos. Olsen has found evidence for spiral chaos

Fig. 2 Time behaviour of the three variables of the simplest screw-type chaotic oscillator. Numerical simulation of the following system of three coupled ordinary differential equations: $\dot{x} = -y$, $\dot{y} = x + 0.55y$, $\dot{z} = 2 - 4z + xz$. Initial conditions: $x(0) = y(0) = z(0) = 1$, $t_{\text{end}} = 152$. For a stereoscopic picture of the same flow see ref. 20.



in the horse-radish peroxidase reaction^{21,22}. Hudson's group, who had done many experiments on the Zhabotinskii reaction in continuous stirred flow conditions²³, also took a second look and found evidence of chaos²⁴. They interpreted their result in terms of the chaotic trajectorial flow described in ref. 4 which corresponds to spiral-type chaos. The present evidence for screw-type chaos (first communicated in ref. 25) does not preclude the possibility of spiral-type chaos in the Zhabotinskii reaction. On the contrary: in abstract systems, the observation of either screw- or spiral-type chaos virtually guarantees that the other type is also possible with slightly changed parameter values²⁰.

We thank Art Winfree for discussions.

Note added in proof: Two-dimensional plots (electrochemical potential against potential of a bromide ion sensitive electrode) obtained in the meantime range from the picture of a fluctuation-triggered monoflop (pseudochaos) to a picture looking like a two-dimensional projection of screw-type chaos.

OTTO E. ROSSLER

*Institute for Physical and Theoretical Chemistry,
University of Tübingen,
and Institute for Theoretical Physics,
University of Stuttgart*

KLAUS WEGMANN

*Institute for Chemical Plant Physiology,
University of Tübingen, 7400 Tübingen, FRG*

Received 26 July; accepted 7 November 1977.

- Rössler O. E. & Hoffmann D. in *Analysis and Simulation of Biochemical Systems* (eds Hemker, M. C. & Hess, B.) 91-101 (North-Holland, Amsterdam, 1972).
- Winfree, A. T. *Science* **175**, 634-636 (1972).
- Rössler, O. E. in *Physics and Mathematics of the Nervous System* (eds Conrad M., Güttinger, W. & Dal Cin, M.) (Springer-Verlag, New York, 1974), *Lecture Notes in Biomathematics* **4**, 546-552 (1974).
- Rössler, O. E. *Z. Naturforsch.* **31a**, 259-264 (1976).
- Khaikin, S. E. *Zh. prikl. fiz.* **7** (1930).
- Adronov, A. A., Khaikin, S. E. & Vitt, A. A. *Theory of Oscillators* 725-729 (Pergamon, New York, 1966).
- Rössler, O. E. *Phys. Lett.* **57a**, 397-398 (1976); *Z. Naturforsch.* **31a**, 1664-1670 (1976); *Bull. Math. Biol.* **39**, 275-289 (1977).
- Noyes, R. M., Field, R. J. & Körös, E. *J. Am. chem. Soc.* **94**, 1394-1395 (1972).
- Field, R. J., Körös, E. & Noyes, R. M. *J. Am. chem. Soc.* **94**, 8649-8664 (1972).
- Juro, J. J. & Noyes, R. M. *J. Am. chem. Soc.* **97**, 5422-5430 (1975).
- Noyes, R. M. & Juro, J. J. *J. Am. chem. Soc.* **97**, 5431-5433 (1975).
- Rössler, O. E. *Z. Naturforsch.* **32a**, 607-613 (1977).
- Li, T. Y. & Yorke, J. A. *Am. Math. Monthly* **82**, 985-992 (1975).
- Junkers, G. thesis, Univ. Aachen (1969).
- Sørensen, P. B. in *Physical Chemistry of Oscillatory Phenomena*, (ed. Gray, P.) Faraday Symposium No. 9, p. 150. (Faraday Society, London, 1974).
- Winfree, A. T. *Science* **181**, 937-939 (1973).
- Rössler, O. E. in *Synergetics - A Workshop* (ed. Haken, H.) 174-183 (Springer, New York and Heidelberg, 1977).
- Kuramoto, Y. & Yamada, T. *Prog. theor. Phys.* **56**, 679-683 (1976).
- Rössler, O. E. *Z. Naturforsch.* **31a**, 1168-1172 (1976).
- Rössler, O. E. in *Synergetics - A Workshop* (ed. Haken, H.) 184-199 (Springer, New York and Heidelberg, 1977).
- Olsen, L. F. thesis, Univ. Odense (1977).
- Olsen, L. F. & Degn, H. *Nature* **267**, 177-178 (1977).
- Graziani, K. R., Hudson, J. L. & Schmitz, R. A. *Chem. Engng J.* **12**, 9-15 (1976).
- Schmitz, R. A., Graziani, K. R. & Hudson, J. L. *J. chem. Phys.* **67**, 3040-3044 (1977).
- Rössler, O. E. *Chemiedozententagung*, Marburg, 1977.

Erratum

In the letter 'Surge activity on the Barnes Ice Cap' by G. Holdsworth, *Nature* **269**, p. 588, in paragraph 5 line 32 for computation¹⁰ read computation¹³; in line 40 for balance¹⁰ read balance¹³. In paragraph 8 line 3 for form⁸ read form¹¹.

Nature Index and Binders

The complete Index for 1976 is available, price £2.50 (UK), US\$5.00 (Rest of World). Copies of the 1975 index are still on sale, price £2.25 (UK), US\$5.00 (Rest of World).

Binders for the journal are also available at £8.00 (UK), US\$16.00 (Rest of World) for three: a year of *Nature* fits into three binders.

All prices include postage. Payment may be made in any currency at the prevailing exchange rate. Orders should be sent, accompanied by remittance, to Macmillan Journals Ltd, Brunel Road, Basingstoke, Hampshire, RG21 2XS, England.

matters arising

The Thera eruption and Late Minoan-IB destruction on Crete

PICHLER and Schiering¹ argue that there is no relationship between the paroxysmal volcanic eruption of Santorini in the Bronze Age and the Late Minoan-IB destructions on Crete, as proposed by Marinatos². The arguments for or against such a relationship are not as clear as they contend. There are several suggestions in their paper which may be disputed and they have been selective in their evidence. Contrary to their major point, the resolution of the pottery chronology is not sufficiently precise to infer a 50-yr gap between the destructions on Crete and the main eruption. The complexities of the relationship between IA and IB pottery styles, recently reviewed by Luce³, have been ignored. There is evidence that the two styles are in part synchronous⁴ and there is considerable doubt as to their stratigraphic distinction³⁻⁵.

Acrotiri was originally destroyed by earthquakes and evacuated before the eruption^{6,7} and was then apparently revisited by a group reconstructing the town, but was never reinhabited⁷. The gap between the earthquakes and the eruption is often quoted as less than two years^{1,3,5}, but this is speculation. The evidence on Santorini does not preclude a series of plausible events such as precursory earthquakes and even small eruptions, which may have discouraged reoccupation for an extended period. We agree with Pichler and Schiering that the eruption was probably very short, but challenge their interpretation of the deposits on Santorini. The occurrence of 3m lithic blocks is not evidence of the onset of caldera collapse, or even circular or polygonal fractures. There are several alternative interpretations, including ejection of lithic blocks during phreatomagmatic explosions, blocks eroded from the vent wall, blocks picked up by mud-flows or pyroclastic flows from the volcano's slopes⁸. The largest lithic blocks on Santorini are predominantly found in deposits interpreted by Bond and Sparks⁹ as mud-flow not ash-flow deposit. Bond and Sparks⁹ also describe late-stage flood deposits and ignimbrites which were emplaced before substantial collapse had begun. A gap of unknown duration between the eruption

and caldera collapse cannot be discounted.

The confidence with which Pichler and Schiering assert that the ash layer was of negligible thickness on Crete is unwarranted. Ninkovich and Heezen⁹ provided firm evidence of the large scale of the eruption and of a high probability of ash-fall on Crete. Their evidence cannot be discarded, as later papers have confirmed the correlation of the Minoan ash layer^{10,11} in RV Vema cores from the Lamont-Doherty Geological Observatory by the presence of the 7,000-yr BP sapropel and chemical analyses. Although the ash layers in cores V10-50 and V10-58 are undoubtedly overthickened by slumping, a reasonable thickness of ash has to accumulate initially to produce a slump. Pichler and Schiering claim to have succeeded in the most difficult of geological tasks, to show that visible remnants of the Minoan ash do not exist on Crete. Their dismissal of the presence of volcanic glass fragments in Cretan soils¹²⁻¹⁴ as due to the use of imported pumice lumps in Minoan households and workshops is unconvincing, and we suggest that this evidence supports ash-fall on Crete. Similarly, to dismiss the possibility of catastrophic tsunamis on the basis of an unpublished theory is unacceptable. Finally we note that the 'isoseismic' data in their Fig. 2 concerning the 1926 earthquake, published over 40 yr ago, seems to be a function of the density of population and hence observation points as well as possible tectonic blocks. We believe, therefore, that more evidence is required to test the theory of Marinatos² relating the Minoan demise to volcanic activity. There

are three time gaps to be considered: that between the earthquake(s) which destroyed Acrotiri and the onset of the eruption, that between the destruction of Acrotiri and devastation of Crete and that between the eruption and caldera collapse, and possible resulting tsunamis. The many uncertainties in the length of these time intervals is such that the volcanic theory cannot be discarded.

R. S. J. SPARKS
H. SIGURDSSON
N. D. WATKINS

Graduate School of Oceanography,
University of Rhode Island,
Kingston,
Rhode Island 02881

1. Pichler, H. & Schiering, W. *Nature* **267**, 819-822 (1977).
2. Marinatos, S. *Antiquity* **13**, 425-439 (1939).
3. Luce, J. V. *Am. J. Archaeol.* **80**, 9-18 (1976).
4. Pendlebury, J. D. S. *The Archaeology of Crete* (Norton, New York, 1900).
5. Page, D. L. *Soc. Promot. Hellenic Studies*, Suppl. Pap. no. 12 (London, 1970).
6. Marinatos, S. (ed.) *Acta int. Sci. Congr. Volcano Thera* (Athens, 1971).
7. Doumas, C. *Antiquity* **48**, 110-115 (1974).
8. Bond, A. & Sparks, R. S. J. *J. geol. Soc. Lond.* **132**, 1-16 (1976).
9. Ninkovich, D. & Heezen, B. C. *Colton Papers* **17**, 413-452 (Butterworths, London, 1965).
10. Richardson, D. & Ninkovich, D. *Geol. Soc. Am. Bull.* **87**, 110-116 (1976).
11. Ninkovich, D. & Heezen, B. C. *Nature* **213**, 582-584 (1969).
12. Boekschoten, G. J. *Opera Botanica* No. 30, 40-48 (Lund, 1971).
13. Vitaliano, C. J. & Vitaliano, D. B. *Am. J. Archaeol.* **78**, 19-24 (1974).
14. Cadogan, G., Harrison, R. K. & Strong, G. E. *Antiquity* **46**, 310-313 (1972).

Possible solar eclipse effect 23 October 1976

LILLEY and Woods¹ operated magnetometers at 10 sites in central and eastern Australia during the solar eclipse of 23 October 1976, and reported a possible eclipse effect at two stations (labelled H and J in their Fig. 1) lying in the path of totality. They observed an event at approximately the time of totality (0638 UT) but pointed out that the event had not moved from west to east at the expected rate. As an eclipse affects the ionosphere over a relatively large area one would not expect the few minutes of totality to produce such a pronounced effect as indicated by their records.

In South Africa the eclipse was partial and of limited duration, starting at sunrise (~ 0420 UT). The records of the magnetic observatories Hermanus (34°S, 19°E) and Grahamstown (33°S, 27°E) should be free of eclipse effects at the time of totality in Australia (0638 UT). The declination variations as recorded at these two observatories over the period illustrated by Lilley and Woods are shown in Fig. 1. It is evident

Matters Arising

Matters Arising is meant as a vehicle for comment and discussion about papers that appear in *Nature*. The originator of a Matters Arising contribution should initially send his manuscript to the author of the original paper and both parties should, wherever possible, agree on what is to be submitted. Neither contribution nor reply (if one is necessary) should be longer than 300 words and the briefest of replies, to the effect that a point is taken, should be considered.

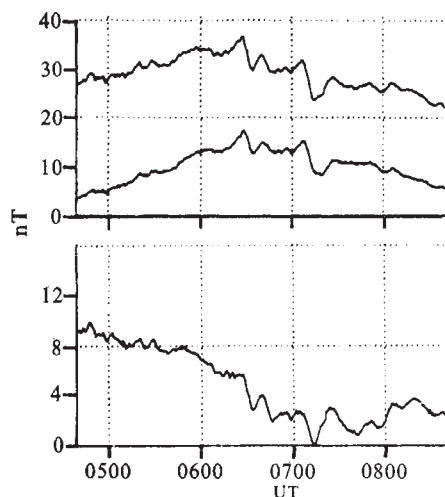


Fig. 1 The declination records for Grahamstown and Hermanus, and the Grahamstown-Hermanus difference curve, plotted on the same scale as used by Lilley and Woods¹. All records are plotted relative to arbitrary zeros.

that the magnetic field in South Africa was slightly disturbed during the interval 0620–0640 UT when the eclipse occurred in Australia. The 0711 event mentioned by Lilley and Woods is also present although the relative amplitude is smaller than in Australia. The Grahamstown-Hermanus difference graph shows a rapid variation at 0628 UT compared to 0638 UT in Australia. The event reported by Lilley and Woods could, therefore, have been a small eastward travelling disturbance. One should, however, be careful not to attribute all sudden changes in the difference curves to a phase shift. High frequency variations at site J are clearly more damped than the corresponding variations at site H. The form and smoothness of the difference curves are in fact determined to a large extent by the relative frequency responses at the two sites, and not by phase shifts, except in the case of the S_q variations. Similar effects can be seen in the Hermanus-Grahamstown comparison. High frequency response at Hermanus is markedly lower than at Grahamstown.

We conclude that the event reported by Lilley and Woods is a manifestation of a minor magnetic disturbance. The variations observed at the Gngangara observatory in Western Australia should give more information on the suggested eastward travelling event.

G. L. M. SCHEEPERS

Magnetic Observatory of the CSIR,
PO Box 32,
Hermanus 7200,
South Africa

LILLEY AND WOODS REPLY—We accept Scheepers' evidence for some widespread minor magnetic activity at the time of the eclipse in south-east Australia. If the source currents for this activity flowed in the overhead ionosphere they (like the S_q currents) may be perturbed by lowered ionospheric conductivity in an eclipse shadow, and we are now examining the

data from our line of instruments for such an effect. We agree with Scheepers that such an ionospheric effect may be difficult to distinguish from local differential induction.

F. E. M. LILLEY
D. V. WOODS

Research School of Earth Sciences,
Australian National University,
Canberra, Australia

Climatic interpretation of $\delta^{18}\text{O}$ and δD in tree rings

THERE has been considerable interest^{1–5} in, and some controversy^{6–7} over the use of the stable isotopes of oxygen and hydrogen in tree rings as measures of past climate. We point out here that variations in isotopic values from the whole ring (or a number of rings) may be in part a function of variations in ring width. As a consequence, whole ring isotope estimates of past climate will not be independent of ring width estimates of past climate. This is a direct result of the relative constancy of either early- or late-wood width in many tree species and of the differences in isotopic composition between early and late wood.

In many species of tree, much of the variation in ring width from year to year is in either the early-wood or the late-wood part of the ring. For instance, in oak⁸ and elm (D. Brett, personal communication) the width of the early wood is approximately constant, while in most conifers the late-wood width is approximately constant⁴. Early wood and late wood generally have different isotopic composition. The cellulose analyses of Epstein and Yapp⁵ and Wilson and Grinstead⁴ both show differences of up to 50‰ in δD . The latter authors show lighter isotopic values in the early wood which they attribute to temperature-dependent fractionation effects. Epstein and Yapp⁵ disagree with this interpretation. Their analyses show that early-wood isotopic values may be either lighter or heavier than those of the late-wood depending on the location of the tree, and their work suggests that late-wood is isotopically lighter (for Douglas Fir) in regions where there is winter snow cover.

Epstein and Yapp⁵ note that "... differences in early- and late-wood δD could cause sampling problems. . .". We elaborate here on one way which such problems might arise, and on their possible implications. Approximate constancy of early- or late-wood width, and isotopic variation across a ring, when considered together, necessitate a relationship between ring width and the isotopic value ($\delta^{18}\text{O}$ or δD) of the whole ring. If E is the early-wood width, L the late-wood and $W (= E+L)$ the total ring width, and if δ_E , δ_L and δ_w are the corresponding mean stable isotope

($\delta^{18}\text{O}$ or δD) values, then

$$\delta_w = \delta_E + (\delta_L - \delta_E) L/W \quad (1)$$

or

$$\delta_w = \delta_L - (\delta_L - \delta_E) E/W \quad (2)$$

These alternative expressions for δ_w arise from the fact that δ_w is a weighted average of δ_E and δ_L . They both contain a term proportional to W^{-1} and thus imply a direct functional dependence of δ_w on W . The sign of the relation between δ_w and W depends on the sign of $\delta_L - \delta_E$ and on whether the tree has relatively constant E or relatively constant L .

As ring width is determined partly by climate (although the relation is often complex⁹), these results imply a dependence of δ_w on climate which would occur even if δ_E and δ_L were constants. Thus, at least a part of any δ_w -climate relationship may be attributable to ring width variations.

Whether this 'ring width effect' contributes significantly to variations in δ_w depends on the magnitude of variations in δ_E , δ_L and either E/W or L/W . For oak, where E is approximately constant and equation (2) is the appropriate expression, Eckstein and Schmidt⁸ give ring width data where E/W varies from 0.23 to 0.63 over the period 1880 to 1969. Since $\delta_L - \delta_E$ may be as high as 50‰ for δD , equation (2) shows that the ring width effect could give δD variations for single whole rings of up to 20‰.

Expected variations in $\delta^{18}\text{O}$ of whole rings due to the ring width effect are difficult to estimate because no suitable early-wood and late-wood isotopic data is available. The values would depend on whether precipitation or temperature-dependent fractionation effects are dominant in controlling tree ring isotopic composition, a subject on which there is considerable disagreement. If precipitation effects dominate then $\delta^{18}\text{O}$ variations of order 2.5‰ could be expected as a result of the ring-width effect. If fractionation effects dominated then even greater variations in $\delta^{18}\text{O}$ could occur.

Gray and Thompson¹ and Libby *et al.*² have found significant correlations between whole ring isotopic values and mean annual temperatures. Their results may be partly due to the ring width effect. Gray and Thompson's¹ data show a $\delta^{18}\text{O}$ range of about 3‰ with lower values corresponding to cooler mean annual temperatures. In their case we expect L to be relatively constant. If $\delta_L - \delta_E < 0$, as suggested by the results of Epstein and Yapp⁵, then δ_w should be lower for narrower rings (see equation (1)), and hence, most probably, for colder years. Since Gray and Thompson use 5-yr means, the magnitude of variations due to the ring-width effect could be as much as 1.5‰.

and could account for a significant part of the variation observed. If $\delta_L - \delta_E > 0$ then the ring-width effect would be in the opposite direction and the variation observed by Gray and Thompson¹ would be even more significant.

The results of Libby *et al.*² are more difficult to interpret. One of their oak samples shows variations in δD and $\delta^{18}O$ of approximately 180‰ and 9‰ respectively (for 30-yr means). This variability is much larger than one would expect from the ring-width effect. As Epstein and Yapp³ point out, however, the magnitudes of the variations observed by Libby *et al.*² are much larger than one would expect from precipitation isotope variations (which Libby *et al.*² state to be the dominant control). Such large variability is even more unexpected when one recalls that Libby *et al.*² use 30-yr means. Nevertheless, their results do correlate reasonably well with central England temperature data.

Isotopic data from tree rings are potentially a powerful indicator of past climate. There is, however, considerable disagreement between the opinions of different workers in this field, and many of the results seem to be mutually incompatible. The simple analysis presented here suggests that variations in whole ring isotope data may be significantly influenced by variations in ring width, and that such isotope data need not be an independent climate indicator. This additional complicating factor may help to explain some of the incompatibilities. We suggest that a comparison between ring width and isotopic variations should be a first step in any analysis of whole ring isotopic data, whether the data comes from whole wood or chemically more specific isotope measurements.

T. M. L. WIGLEY
B. M. GRAY
P. M. KELLY

Climatic Research Unit,
University of East Anglia,
Norwich, UK

1. Gray, J. & Thompson, P. *Nature* 262, 481-482 (1976).
2. Libby, L. M. *et al.* *Nature* 261, 284-288 (1976).
3. Libby, L. M. & Pandolfi, L. J. *Nature* 266, 415-417 (1977).
4. Wilson, A. T. & Grinsted, M. J. *Nature* 257, 387-388 (1975).
5. Epstein, S. & Yapp, C. J. *Earth planet. Sci. Lett.* 30, 252-266 (1976).
6. Epstein, S. & Yapp, C. J. *Nature* 266, 477-478 (1977); reply by Libby, L. M. & Pandolfi, L. J. 266, 478 (1977).
7. Perry, D. A. *Nature* 266, 476-477 (1977).
8. Eckstein, D. & Schmidt, B. *Angew. Botanik* 48, 371-383 (1974).
9. Fritts, H. C. *Tree Rings and Climate* (Academic, London, 1976).

WILSON REPLIES—Wigley, Gray and Kelly make the valid point that isotopic measurements on whole wood may give results which merely reflect the ring width. One aspect of the problem was discussed by Wilson and Grinsted¹ who pointed out that whole wood contains a mixture of lignin and cellulose which have different isotope compositions. Since

early and late wood can have different lignin to cellulose ratios, isotopic measurements on whole wood may merely reflect the ratio of late wood to early wood which may in turn be a simple reflection of ring width. Wilson and Grinsted suggested that this and other problems could be avoided if isotope dendroclimatologists measured only pure cellulose or pure lignin.

The comments of Wigley *et al.* on the climatic interpretation of the isotopic measurements on tree rings and the relationship of these to past climate involve the fundamental principles on which isotope dendroclimatology is based. Much of the controversy in this field stems from certain misconceptions of the aims of isotope dendroclimatology and the basic plant physiology involved.

Isotopic dendroclimatology studies face two problems. First, at what times of the year is the material that is ultimately formed into wood components fixed from the atmosphere? And second, how does the isotopic composition of a wood component such as cellulose vary with changing climate?

A conifer manufactures photosynthates at all times of the year except when climatic conditions are unsuitable, due for example to low temperatures or drought stress. These photosynthates are stored for short or extended periods before they are laid down as wood. Most conifers (*Pinus radiata* is one of the exceptions) lay down wood only during a brief period of the year. The actual period of wood deposition is not controlled by net photosynthesis but is under hormone control, the production of hormones being controlled principally by day length and to a lesser extent by temperature. For example, in trees suffering from drought stress wood can be laid down during periods when net photosynthesis is negative².

The above considerations apply to any isotopic work on trees. But in the case of isotopic work on oxygen and hydrogen another factor is important. The isotopic composition of rain or snow depends on many factors including the temperature history of the air masses which bring the precipitation to an area³. This is not the end of the problem, however. Once the water is taken into the tree transpiration processes in the leaves can cause very large fractionations particularly in arid environments⁴.

Wigley *et al.* state that "early wood and late wood generally have different isotopic composition". Leaving aside the point, discussed above, that early and late wood generally have different D/H ratios because they are a variable mixture of more than one compound each with its own isotopic composition. Late-wood cellulose and early-wood cellulose generally have different isotope composition because they represent material fixed from the atmosphere at different

times of the year and hence in different climatic conditions. For example in New Zealand *Pinus radiata* early wood is laid down in the spring and summer and late wood in the winter, that is, early wood is laid down at a warmer time than the late wood. On Mt Lemon, Arizona, *Pinus ponderosa* lays down early wood in the spring and late wood in the summer⁵. That is, early wood is laid down at the colder time than the late wood. The important point is that the isotopic composition is determined by the climate at the time the carbohydrate was fixed from the atmosphere and not whether it is early or late wood that is being formed.

Isotope dendroclimatology can never be expected to produce a mean annual temperature curve as might be produced by a meteorological observer or from measurements of speleothems⁵. It will only produce a curve representative of some period of the year for example 'a spring and early summer' temperature curve. Different trees may fix from the atmosphere the carbohydrate ultimately formed into wood at different times of the year and hence record climate at different periods of the year. Obviously a deciduous stress such as an oak growing in New Zealand only carries out photosynthesis in the spring and summer, whereas at the same site a conifer such as *Pinus radiata* can fix material from the atmosphere all times of the year. Thus the wealth of palaeoclimatic data recorded in the isotopic ratios of the constituents of tree rings and the superb time base provided by dendrochronology makes the potential of isotopic dendroclimatology very great indeed.

A. T. WILSON

Chemistry Department,
University of Waikato,
Hamilton, New Zealand

1. Wilson, A. T. & Grinsted, M. J. *Nature* 265, 133-135 (1977).
2. Fritts, H. C. *Tree Rings and Climate* (Academic, London, 1976).
3. Friedman, I., Redfield, A. C., Schoen, B. & Harris, J. *Rev. Geophys.* 2, 177-224 (1964).
4. Wershaw, R. L., Friedman, I. & Heller, S. J. in *Advances in Organic Geochemistry* (eds Hobson, G. D. & Speers, G. C.) 55-67 (Pergamon, Oxford, 1966).
5. Hendy, C. H. & Wilson, A. T. *Nature* 219, 48-51 (1968).

GRAY AND THOMPSON REPLY—The 'ring width effect' discussed by Wigley *et al.* is relevant in two regards to our findings of significant correlation between $\delta^{18}O$ values in cellulose and mean annual temperatures¹. First, as stated by Wigley *et al.*, there is a possibility that an appreciable fraction of the variation in the $\delta^{18}O$ values may be due to varying proportions of early and late wood in the tree rings used for analysis. Second, the possibility exists of a sampling error being introduced when a 5-yr group of rings is analysed.

To evaluate the first effect we have made measurements of late-wood width (*L*), early-wood width (*E*), and total ring

width (W) for the entire section of tree used for production of the climate curve¹. Values of L/W ranged from 0.05 for the largest value of W to 0.32 for the smallest W . Further, $^{18}\text{O}/^{16}\text{O}$ isotopic analyses were carried out on early and late wood from a number of selected rings from the tree. (These were necessarily rings having larger values for W due to difficulties associated with the sampling of narrow rings and insufficiency of material for analysis.) The range of values found for the difference in isotopic composition between early and late wood ($\delta E - \delta L$) was found to be 0.5 to 0.8‰.

Thus using equation (1) as stated by Wigley *et al.* with values of $\delta L - \delta E = -0.8$ ‰ and $L/W = 0.32$ the maximum variation in δW due to the term dependent on total ring width is found to be 0.26‰. Thus the maximum contribution to the variation is estimated at somewhat less than 10% of the total variation observed (3‰) in a 100-yr-old tree. Furthermore, by taking 5-yr groups of rings this effect will tend to be minimised since the tree in question contained few narrow rings and the effect is much less significant for wider rings (L/W of the order 0.05). In general, therefore, we expect the effect to be considerably less than the 10% suggested above.

The measured values of tree ring width (W) for the Edmonton spruce, after growth curve corrections were made, show little or no correlation with mean annual temperatures. We conclude that in this case, the contribution of ring-width effects to variation in $\delta^{18}\text{O}$ of the cellulose is minimal. This finding, however, does not detract from the potential significance of the effect for other trees in other climate zones. Care should be taken to select trees which are 'complacent', that is, those with ring widths showing little annual variation. (This is in contrast to the requirements of those establishing climate curves from ring-width measurements².) The second possible effect of ring-width variation, that of introduction of a sampling error, arises because, when sampling 5-yr groups of rings³, no attempt was made to normalise the amount of wood contributed by each of the rings to the material subsequently analysed. When calculating mean annual temperatures for each 5-yr period, however, an arithmetic mean was used, implying equal contribution from each ring and hence implying equal ring width throughout the 5-yr group. To estimate the error introduced by this procedure, we measured total ring widths for the entire tree. Using mean annual temperature (T) from climate records we calculated the expected isotopic composition of the cellulose in each ring using the relation $\delta^{18}\text{O} = 1.3T + 20.5$ (‰) (ref. 1). The mean $\delta^{18}\text{O}$ for each 5-yr group of rings was then calculated first using

Table 1 Comparison of ring-width weighted mean, non-weighted mean and measured $\delta^{18}\text{O}$ values for cellulose extracted from 5-yr groups of tree rings

Period	$\delta^{18}\text{O}$ (‰) weighted mean	$\delta^{18}\text{O}$ (‰) unweighted mean	$\delta^{18}\text{O}$ (‰) measured
1894-99	23.4	23.5	23.4
1900-04	24.2	24.4	24.1
1905-09	24.7	24.6	24.7
1910-14	24.6	24.6	24.7
1915-19	24.0	24.0	24.1
1920-24	23.8	24.0	24.2
1925-29	23.7	23.5	23.6
1930-34	24.6	24.7	24.7
1935-39	23.2	23.3	22.9
1940-44	24.9	24.6	24.5
1945-49	23.8	23.7	23.3
1950-54	23.7	23.3	23.3
1955-59	24.4	24.5	24.5
1960-64	25.1	25.3	25.2
1965-68	23.5	23.7	24.0

the ring-width data as a means of weighting the $\delta^{18}\text{O}$ contribution of each ring and second, assuming equal contributions for each ring in a given 5-yr group. The calculated weighted and unweighted means, together with measured values are shown in Table 1. The largest difference between weighted and unweighted calculated value is 0.4‰, approximately twice the estimated precision of the measured data (± 0.2 ‰). It seems that errors introduced elsewhere in the procedures used, are usually equal to or greater than those due to sampling errors. It should be pointed out, however, that trees showing marked variability of ring width with climate may well present sampling problems of this kind.

J. GRAY
P. THOMPSON

Department of Physics,
University of Alberta,
Edmonton, Alberta,
T6G 2E1, Canada

1. Gray, J. & Thompson, P. *Nature* **262**, 481-482 (1976).
2. Fritts, H. C. *Tree Rings and Climate* (Academic, London, 1976).

LIBBY AND PANDOLFI REPLY—Pandolfi has shown that isotope variations in tree rings are correlated with climate variations local to the trees. Wigley *et al.* speculate that isotope variations in tree rings may be correlated with ring widths. Then if they are correct, it must follow that ring widths are correlated with climate.

We consider that the master chronologies of ring widths from the European tree laboratories will not correlate with local climate variations. We found no correlation between isotope measurements for a German oak and B. Huber's master oak ring chronology.

LEONA MARSHALL LIBBY
LOUIS J. PANDOLFI

Environmental Science and Engineering
Department,
University of California, Los Angeles,
California 90024

Amide nitrogen is unlikely to be a proton acceptor

In their article on the mechanism of catalysis of acid proteases and extended to proteases generally, James *et al.*¹ make several statements to the effect that a tyrosine donates its proton to the amide nitrogen of the scissile bond in a peptide substrate. This mechanistic idea is similar to that suggested for hydrolysis of peptide substrates by carboxypeptidase^{2,3}. But the peptide nitrogen is not the basic site in an amide bond and cannot accept a proton, a hydrogen bond, or a metal ion in the ground state of an enzyme-substrate complex⁴. This restriction holds with even more force when the amide carbonyl oxygen interacts with a proton or a metal ion, as is thought to be the case with carboxypeptidase.

There are many data to support the view that the basic site in an amide bond is the carbonyl oxygen so that either a proton^{5,6} or a metal ion^{7,8} associates at that atom. These conclusions are supported by theoretical calculations which indicate a strong preference for protonation of amides at the carbonyl oxygen^{9,10}. Unless there is enormous strain associated with substrate binding, the earliest stage in a hydrolysis mechanism during which an amide nitrogen may become a proton, hydrogen bond, or metal ion acceptor is concerted with nucleophilic attack at the amide carbon which results in both the carbon and nitrogen of the amide bond taking on tetrahedral character⁴. At this point the nitrogen has lost its amide character and has become an amine nitrogen. This principle seems to have been accepted in a recent discussion of the mechanism of thermolysin which is thought to be similar to that of carboxypeptidase¹¹. Though the authors of the acid protease paper¹ may feel that they did not violate this principle in their own minds, their statements coupled with proposed mechanisms in the literature cited are apt to mislead many readers.

R. BRUCE MARTIN

Chemistry Department,
University of Virginia,
Charlottesville, Virginia 22901

1. James, M. N. G., Hsu, I-N. & Delbaere, T. J. *Nature* **267**, 808-813 (1977).
2. Quicho, F. A. & Lipscomb, W. N. *Adv. Protein Chem.* **25**, 1-78 (1971).
3. Lipscomb, W. N. *Accounts Chem. Res.* **3**, 81-89 (1970); *Tetrahedron* **30**, 1725-1732 (1974); *Proc. Nat. Acad. Sci. U.S.A.* **70**, 3797-3801 (1973).
4. Sundberg, R. J. & Martin, R. B. *Chem. Rev.* **74**, 471-517 (1974).
5. Martin, R. B. *J. Chem. Soc. Chem. Commun.* 793-794 (1972).
6. Martin, R. B. & Hutton, W. C. *J. Am. Chem. Soc.* **95**, 4752-4754 (1973).
7. Morris, P. J. & Martin, R. B. *Inorg. Chem.* **10**, 964-968 (1971).
8. Freeman, H. C. *Adv. Protein Chem.* **22**, 257-424 (1967).
9. Bonaccorsi, R., Pullman, A., Scrocco, E. & Tomasi, J. *Chem. Phys. Lett.* **12**, 622-624 (1972).
10. Pullman, A. *Chem. Phys. Lett.* **20**, 29-32 (1973).
11. Kester, W. R. & Matthews, B. W. *Biochemistry* **16**, 2506-2516 (1977).

reviews

American Natural History

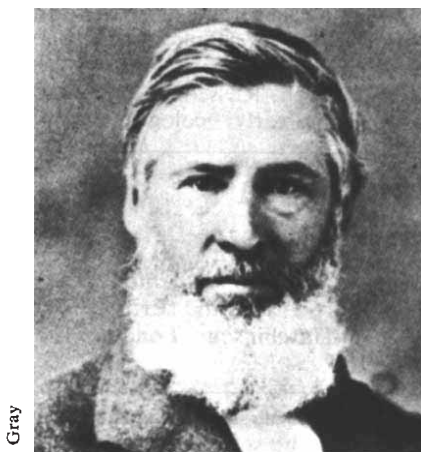
Jack Major

Natural History in America. By Wayne Hanley. Pp.xiii+334. (Harper and Row: London; Quadrangle/New York Times Books: New York, 1977.) £8.95; Fl.40.65; \$14.95.

THE author compiles and discusses excerpts from the record of natural history observations during settlement of the US over the past two centuries. Ornithology is emphasised. Colour plates of paintings made by the eighteenth- to nineteenth-century naturalists are included. Unfortunately, the book has no index, nor reference list. Hanley's method is to quote passages from the early naturalists' published work and to relate them to then contemporary events in the biological sciences at large. Thus, Linnaeus and his emissary Peter Kalm, as well as Darwin's work, are very much present in early American natural history. The author introduces his book modestly: "This is not a profound volume. It is essentially a collection of readable passages selected from the writings of great American naturalists" (page xi). Too modestly. The events and thoughts described and quoted have acquired a profundity in our time when America has used up so much of its bountiful inheritance and when so many are so concerned that some of our wonderful natural heritage be passed on to future generations.

Modern naturalists will be deeply impressed, also, by young Meriwether Lewis setting forth on his extended (1804-1806) field trip into the wilderness bearing a hand-written, open letter of credit from President Jefferson authorising unlimited funds on the credit of the US Government. The Indian woman, Sacajawea, who guided him through the Rocky Mountains was of more help.

The book opens with a chapter, "In the Beginning: Catesby", on "an exception to the general rule of incompetence or worse among early eighteenth-century naturalists"—Mark Catesby's hand-coloured, illustrated *Natural History of Carolina, Florida and the Bahama Islands of 1730-1748*. His stiff but accurate paintings are followed by discussion of John and William Bartram, collectors of plants and ornithologists; of Alexander Wilson, with



Gray

his colour-illustrated, nine-volume *American Ornithology*; and climaxed by Audobon's magnificent paintings and vivid writings and his collaboration with the mammalogist John Bachman. The western part of the present US was almost unknown to these naturalists, and one gets the impression they regarded its biological riches from the standpoint of the Boston dowager who described her route to San Francisco, over 4,000 km distance, by way of Newton, a suburb of Boston. But they were fully occupied with the area east of the Mississippi River from Labrador south to Florida, and the modern naturalist can only appreciate and envy their discoveries.

Lewis and Clark explored westward from St Louis, Missouri, to the Pacific. They are given a chapter. Hanley emphasises their written interest in the grizzly bear and Indian horses, almost to the exclusion of the other myriad new forms of plant and animal life they encountered on their epic journey. Horses are of obvious interest to anyone on a walk from St Louis to the mouth of the Columbia River. And anyone afoot, armed or unarmed, who has seen a grizzly moving ponderously but so fluidly about its own business and then throw up its head to stare at the intruder, will sympathise with Lewis and Clark's deep interest.

Although Hanley's book has in its title *America*, Canadian contributions are absent. The voyageurs who performed miracles of physical labour and daring in penetrating America toward the Pacific, and the Hudson Bay and North-West Company factors who em-



Agassiz

ployed them seem to have written little on the Natural History of these virgin lands, for reasons of illiteracy or business priorities. But Sir Alexander Mackenzie's journals describing his explorations from Lake Athabasca to the Arctic Ocean (1789) and then to the Pacific at Bella Coola (1792-1793) are not devoid of Natural History observations. True, he noted birch and spruce because they could furnish bark and gum to repair his canoe, various shrubs because their berries were edible, devils club (*Oplopanax horridum*) because it hindered his progress; but transportation and food of necessity were his primary interests. And he did hesitate on his perilous return journey to measure one of the giant cedars (*Thuja plicata*) which he noted characterised the magnificent coniferous forest of the Pacific slope.

But Hanley has plenty of material in the US alone. Thomas Nuttall collected plants throughout much of the US, wrote a catalogue of American plants, a botanical textbook, a bird manual, spent 11 years at Harvard University, and ended up on an inherited estate in England. Thomas Say, on the other hand, seems to have been totally disinterested in money, but he was interested in collecting, travelling, and he advanced Natural History knowledge, uniquely in entomology and conchology.

N. M. Hentz, too, was eccentrically interested in less well-known organisms, spiders in his case. The climax comes with Rafinesque. It was a taxonomic climax, only exceeded by the collections of the exploring expedition

around the world under Lieutenant Charles Wilkes, who not only went to Antarctica but whose shore parties contributed to our knowledge of Mount Shasta in northern California. Hanley also devotes a chapter to the whalers, including the novelist Herman Melville who was a crew member in the bloody business "until he jumped ship at Marquesas Island in July, 1842, preferring to live among cannibals".

And finally comes the first writer and naturalist in Hanley's collection who looked at nature as a whole, with sympathy. Did Thoreau's words exert much influence? Probably not, but we are discovering them anew. He is a charming prophet.

In the following chapter, Agassiz and Asa Gray differed over Darwin's ideas without adding much to them, and they gave Harvard an early taste of dissension among its faculty.

Academic nastiness accompanies Cope and Marsh in their disputes over priority of publication, ownership of specimens, money, and anything else handy one feels. Living reptiles and amphibians found a specialist in J. E. Holbrook.

Elliott Coues was an improbable cavalryman, briefly a member of the mystical Theosophical Society of India, a superb ornithologist whose description and advocacy of bird identification using a shotgun is illuminating. Hanley progresses from the "I do not protect birds, I kill them" school of ornithology to preservationist attempts to stop some of the slaughter of wildlife. First, to stop commercial use of bird skins to decorate women and, later, to save enough big game so hunting could continue. Protagonists of the latter crusade included W. T. Hornaday and Theodore Roosevelt. In the western US, John Muir took up some of the same themes Thoreau had broached in the east. He, too, was ahead of his time in suggesting that not every last resource should be used up by the generation then in propertied possession.

Popularisers of Natural History such as John Burroughs and Ernest Thompson Seton aroused interest in the life histories of mammals in general. If Seton was an egotist, he also had much field experience and was a competent mammalogist. Burroughs deliberately kept the observer (himself) in the forefront of his observations and descriptions, to the detriment of both. Louis Agassiz Fuertes seems to have been a paragon, not only as a painter of birds but as a fine field companion.

An appreciation of the place of predators in the scheme of nature began to appear only with C. Hart Merriam, working within the government. Aldo Leopold progressed to this view, wrote

a textbook which made wildlife management a field of study, and suggested a philosophy of land use in consonance with Thoreau's ideas. But their ideas are still ahead of the economic and commercial mores of our time. Our time is one of continuing pollution of our environment, and Rachel Carson has the last chilling word.

Each person will have his own list of people, quotes, ideas, landmarks, interactions left out of Hanley's little book. The early geologists such as Gilbert, Dutton, Powell, Russell in the western US, and the early ecologists such as

Forbes, Clements, Cowles, Harper, Kearney, Sears, Weaver would be mine. But even if the scope of American Natural History over the past 200 years seems occasionally to approach a Boston ornithologist-dowager's view, Hanley has produced a wonderfully readable and useful account. I hope it will stimulate readers to learn more of the background of our present ecological crises and the science that works toward their solution. □

Jack Major is Professor of Botany at the University of California at Davis.

Running a bird reserve

Minsmere: Portrait of a Bird Reserve. By Herbert Axell and Eric Hosking. Pp. 256. (Hutchinson: London, 1977.) £7.50.

BASICALLY, as its title suggests, this is the story of the setting up and running of a reserve. It is also, however, rather more than that. Minsmere, on the Suffolk coast, is one of the largest and most famous reserves run by the Royal Society for the Protection of Birds (RSPB), and the author was the warden of the reserve for 17 years and was responsible for many of the innovations carried out there.

Before the last war, Minsmere was a low-lying area of farmland, vulnerable to flooding by the sea. In June 1940, it was intentionally flooded as part of Britain's coastal defences against possible invasion by the Germans. In short, the birds approved and by the end of the war a number of rare species were established there. In 1947, the RSPB acquired a lease of the new marshes, together with adjacent areas of heath and woodland.

The book describes the birds of all the Minsmere habitats, giving particular prominence to those that are rare elsewhere in Britain. There are four of these; three—the Bittern, Marsh Harrier and Bearded Tit—are reed-bed birds. Once common over wide areas of Britain, these species have slowly dwindled with the general drainage, particularly of the Fens. The Marsh Harrier and Bittern are reduced to very small numbers in the UK at present, indeed at one stage Marsh Harriers bred nowhere apart from Minsmere. It is one of those wry complexities which nature all too frequently sets for conservationists, that anxious parent Bitterns occasionally satisfy the needs of their clamouring young with baby Marsh Harriers from an adjacent nest. The Bearded Tit is more numerous than the other species and at the

moment more widespread. It was, however, once scarcer and at that time Minsmere was the essential base from which, at a later date, the other colonists spread out.

Maintaining habitat in the stage at which the birds need it, is not just a matter of sitting about; it needs active management. This is especially the case with the fast growing reed-beds. The author not only discusses his day-to-day problems, but also describes new developments which he initiated. Over a series of winters (so the breeding birds would not be disturbed), bulldozers gradually made a series of shallow pools, now famous as "The Scrape". To these, and their gravel-covered islands, birds came in large numbers. Among these were the terns, whose nesting areas along the beach were almost ceaselessly disturbed by people; and waders, whose shallow feeding areas are now so scarce. The other bird to approve of these moves was the Avocet—the fourth species to merit a chapter to itself. With only one other nesting colony in Britain (on Havergate, another RSPB reserve), this was a welcome development.

Perhaps one of the points that emerges from this book is how successful a man-made marsh can be. One cannot, at will, reproduce a mature oakwood, but one can reproduce quite quickly many of the necessary features of a marsh. As drainage relentlessly goes on, perhaps more organisations could be persuaded to find low-lying land which could be used to make amends, albeit in a small way, for some of the losses.

The book is beautifully and profusely illustrated, mostly by Eric Hosking's own photographs. My only criticism is that some of the colour photographs are reproduced at six to a page, and the small size does not permit one to fully appreciate their quality.

C. M. Perrins

C. M. Perrins is Director of the Edward Grey Institute of Field Ornithology, University of Oxford, UK.

Butterfly origin

The British Butterflies: Their Origin and Establishment. By R. L. H. Dennis. Pp. xviii + 318. (E. W. Classey: Faringdon, UK, 1977.) £10.

THE message of this book is simple: recent work on the British Pleistocene suggests that all Britain's butterflies must have been exterminated by cold before the amelioration of climate which began 10–15,000 yr ago; no existing populations of butterflies in the British Isles, including various endemic "subspecies", can be older than that. This view contrasts with the earlier ideas of B. P. Beirne and E. B. Ford, who, working with a radically different and now outmoded reconstruction of Pleistocene events, considered that at least some of our butterflies had existed continuously in Britain for 100,000 yr or more. Dennis's book has some subsidiary themes, one being that phenotypic divergence of a butterfly population bears no direct relationship to either its antiquity or total genetic divergence. I agree with both of these conclusions.

Dennis's treatment of the subject is based on a wide reading of the work on the Pleistocene, local literature pertaining to the distribution and ecology of British butterflies, and an eclectic selection of genetical texts. Two extremely relevant areas of work are completely ignored: MacArthur–Wilson island biogeography theory, aspects of which Dennis apparently rediscovers for himself, and Paul Ehrlich's school of work on *Euphydryas* colony structure. He does refer once to the classic paper of Ehrlich and Raven on butterfly and plant coevolution. In doing so, he misspells Erlich (*sic*), gives the wrong publication date, and apparently misunderstands the particular ideas of biochemical coevolution involved. Further, in discussing phenotypic divergence, antiquity and evolutionary rates, Dennis makes no reference to work on rates of protein evolution.

If the essential message of the book is simple, and key areas of the literature have been ignored, why does Dennis's main text span 270 close-set pages? Because he frequently uses whole sentences where one word (or even none!) would suffice. He is at pains at every point to indicate that he has thought of every possibility. He attacks the same problems again and again, using the very few lines of evidence open to him. As a result, instead of being a delightful natural history detective story, the book is bewildering. Alternative hypotheses are not presented and dealt with in a logical way, step by step. Instead, we are fed with an unending torrent of jargon-laden fact, theory and specula-

tion, leaving the reader in a state of dizzying confusion. Beneath this, many of Dennis's ideas are sound, and one gets an impression of tremendous enthusiasm. If so, it is unfortunate that the book has been published in its present, clearly unedited state, because both the author and his intended audience are the losers.

In short, although it will have some value as a reference source, and for Dennis's original contributions on the analysis of distribution, I cannot recommend this as a good book. I looked forward with keen anticipation to reading it, which is perhaps partly why I find it disappointing. But it is the language which really spoils it. For example, on p243 we are told that "Bearing in mind the reticulate control pathways of atmospheric inputs

into local ecosystems, and the original and ultimate dependence of contemporary environments on atmospheric influences, though modified by the interaction of living and non-living components, and thus, in the attempt, to avoid recourse to simplistic determinism, it is suggested that the increasing assumption of oceanic conditions to the north and west of Britain has resulted in a primary structure of selection gradients." Such gems of egregious obfuscation abound throughout the whole work. Only the dedicated will probe below this forbidding mask.

R. I. Vane-Wright

R. I. Vane-Wright is Head of the Butterfly Section at the British Museum (Natural History).

Chemical dynamics

Dynamics of Molecular Collisions. Parts A and B. *Modern Theoretical Chemistry.* Vol. 1. Pp. 318. \$39.50. Vol. 2. Pp. 380. \$47.40. Edited by W. H. Miller. (Plenum: New York and London, 1977).

IN the past fifteen years, there has been much progress in the field now known as chemical dynamics. The development of the crossed-molecular-beam technique provided a means of studying chemical reactions in more detail, and many new and powerful experimental techniques are now being used in this area. It is now reasonable to ask an ultimate question in kinetics: what is the cross-section for scattering at a particular angle from each state i of species A to every state f of a product B at any relative energy? This experimental activity has led chemists into scattering theory; previously, this had been a province of physics. New theoretical techniques and models have been developed by chemists for handling molecular collisions, and some of these are now finding applications in certain areas of physics.

These first two volumes in the series *Modern Theoretical Chemistry* contain thirteen chapters, each by a different author, on theoretical techniques that are appropriate for describing chemical dynamics. The first, by W. A. Lester, Jr, is on the N coupled-channel problem. The author likens coupled-channel methods to *ab initio* computations of molecular properties, for they provide a source of reliable data in cases where experiments may be difficult and they serve as a test of approximate methods. The second, by H. Rabitz, is devoted to effective Hamiltonians; it describes various means of reducing the number of rotational

channels that need to be considered explicitly. There is a chapter by D. A. Micha on the use of optical potentials in molecular collisions—their name comes from the analogy between the wave-mechanical aspects of molecular collisions and electromagnetic propagation in an inhomogeneous medium with a complex index of refraction. There is a chapter on vibrational energy transfer by H. K. Shin, and one on the scattering of atoms and molecules from solid surfaces by G. Wolken, Jr. The final chapter in Part A is a very clear account of the theory of non-radiative processes by W. Siebrand.

Part B contains articles on classical and semi-classical methods. It begins with a chapter on classical trajectory methods by R. N. Porter and L. M. Raff. The second chapter, by P. J. Kuntz, is an excellent account of the effect of features of the potential energy surface on molecular collisions. The third chapter, by W. L. Hase, is concerned with the dynamics of unimolecular reactions, and in particular with the RRKM theory. M. S. Child contributes a useful chapter on semi-classical methods and J. C. Tully writes on non-adiabatic processes in molecular collisions. The sixth chapter, by P. Pechukas, is on statistical approximations in collision theory, that is, on the transition-state theory. The final chapter, by R. D. Levine and R. B. Bernstein, is on the thermodynamic approach to collisions; it makes use of information theory.

The topics have been well selected and the authors are known for their work in the field. These two volumes will be very useful to research workers entering the field.

A. D. Buckingham

A. D. Buckingham is Professor of Chemistry at the University of Cambridge, UK.

Mechanisms of animal locomotion

Mechanics and Energetics of Animal Locomotion. Edited by R. McN. Alexander and G. Goldspink. Pp. 346 (Chapman and Hall: London, 1977.) £15.

FOR nearly 30 years from the early 1930s, Sir James Gray and his colleagues at Cambridge published a steady stream of papers analysing how animals move in locomotion. There were distinguished earlier and contemporary contributions by others, but Gray set the foundation for the modern development of the subject. His body of work was largely concerned with the kinetics of locomotion and explained how the movements of most groups of animals exerted appropriate thrusts against the environment and produced locomotion. Some progress was made with explaining the nervous coordination of locomotory movements in a few animals and the energetics of locomotion, but there was little advance in the formidably difficult subjects of the aerodynamics or hydrodynamics of animal movement through fluids. During this period, studies of animal locomotion had their firmest links with comparative anatomy, exemplified in the magnificent and detailed functional analysis of the locomotory musculature of arthropods by Sydney Manton (whose book on arthropod locomotion will be published later this year). The physiological dimensions remained relatively underdeveloped alongside this.

The past 20 years have seen a remarkable change in the subject. Previously, animal locomotion was a specialised, rather isolated topic which could be safely ignored by most zoologists. Now, it has become a coherent, mainstream subject drawing together a great variety of scientific knowledge and expertise. As the preface to this volume says with some justice: "it seems a good time to produce a book".

This is an exceptionally good one. It deals with the entire subject of animal locomotion from muscle physiology and nervous coordination to energetics and the kinetics and dynamics of locomotion. Nearly all the topics discussed are difficult for the non-specialist to penetrate, but here they are introduced in an unusually clear, economical and intelligible style. Every concession is made to the reader, and an undergraduate with no previous knowledge of any of these subjects could read this book without much difficulty and with enormous profit. Inevitably, this has led to some simplification and loss of

detail. There is a good chapter on coordination of invertebrate locomotion by Delcomyn and another, shorter, chapter on crawling and burrowing in soft-bodied invertebrates by Trueman and Jones. Otherwise, invertebrates (arthropods in particular) get rather brief treatment and the main emphasis is on the locomotion of vertebrates. Viewed as a whole, however, the book provides an excellent account of the current state of our understanding of the mechanisms of animal locomotion.

The book starts with chapters on the design, mechanics and energetics of muscles involved in locomotion, the coordination of invertebrate and vertebrate locomotion, the energy cost, and then concludes with chapters on terrestrial locomotion, crawling and bur-

rowing, swimming, flying and, finally, the locomotion of Protozoa and single cells. There are good bibliographies at the end of each chapter and the book is well illustrated and produced. The editors are to be congratulated on having recruited a very authoritative team of nine authors and, still more, for having achieved a remarkable consistency of style and uniformity of treatment throughout the book. This is not a compendium of separate, contributed articles as so often happens with multi-author volumes; it has rather been welded into a coherent book.

R. B. Clark

R. B. Clark is Professor of Zoology at the University of Newcastle upon Tyne, UK.

Cell division

Mechanisms and Control of Cell Division. Edited by Thomas L. Rost and Ernest M. Gifford. Pp. vii+387 (Dowden, Hutchinson and Ross: Stroudsburg, Pennsylvania; Halsted/Wiley: New York and Chichester, UK, 1977.) \$31.50; £18.75.

BIOLOGY is suffering a proliferation of books, many of which are simply new combinations of material that has already seen printer's ink at least one and often more times. Sometimes this is justified; a comprehensive well-planned collection can usefully summarise and occasionally illuminate a particular field. This expensive little collection, though containing some new material, does neither. The editors describe it as a "compendium of several areas of cell cycle research, especially those dealing with plants. Some . . . reviews . . . new ideas and the results of recent research" for "researchers and advanced students interested in cell behaviour". The precise purpose is not stated. Was it perhaps simply to market another book? Certainly, the title has all the keywords for the big sell.

In the chapter summarising their elegant dissections of the cell cycle in *Allium* meristems, Giménez-Martin *et al.* point out that in the strict sense "cell division" refers only to mitosis and cytokinesis but is often loosely applied to all processes of proliferation including interphase growth. In making their selection, the editors seem to have been confused about which sense to follow. As a result the collection, which contains some good material, is a hotch-potch.

Gurley *et al.* review their work on biochemical events associated with proliferation of an animal cell-line in culture, Douvas and Bonner re-describe the isolation of contractile proteins from liver chromatin. There follow two interesting reviews of plant hormones and cell proliferation, a list of factors affecting plant cell cycles, a catalogue of structural changes in plant and animal nuclei during replication, two reports on plant cell ultrastructure during mitosis and one of several current hypotheses purporting to explain chromosome movement. The penultimate article is part 1 of an anatomy of cell division in euglenoid algae and the ultimate a detailed description of *all* the stages of mitosis and meiosis and their variations in basidiomycotinal fungi. Several contributions have the prose density of a Patrick White novel.

Despite its dynamic title the collection contains much that is purely descriptive. Being diverse, it lacks the connecting thread that makes a cover-to-cover reading easy or inspiring. Being thinly spread, it fails as a definitive reference source.

The book is strongly bound, but paperbacking would have sufficed and might have halved the price. There is a subject but no author index, nor notes on contributors other than the editors. Many of the numerous electron-micrographs are reproduced with a fuzziness that seriously obscures fine detail. The layout is wasteful: a line diagram of six microtubules occupies a page, as do many other diagrams originally published at a quarter of the size.

L. Martin

L. Martin is a member of the Hormone Physiology Department at the Imperial Cancer Research Fund Laboratories, London, UK.